Algèbre (bi)linéaire, espaces euclidiens

Fabien Priziac

Licence 3 Spécialité Mathématiques, année universitaire 2022-2023

Table des matières

Index des notations

Ci-dessous, $\mathbb{K}$ désigne un corps commutatif quelconque. E et F désignent deux espaces vectoriels sur $\mathbb{K}$. f désigne une application linéaire de E dans F et g désigne un endomorphisme de E. $v, v_1, \ldots, v_k$ désignent des vecteurs de E. $\mathcal{B}$, $\mathcal{B}'$ désignent deux bases de E. A désigne une matrice à coefficients dans $\mathbb{K}$. $\mathcal{C}$ désigne une base de F. n et p désignent des entiers naturels non nuls. i et j désignent deux nombres de l'ensemble $\{1, \ldots, n\}$. M désigne une matrice carrée de taille n. P désigne un polynôme de $\mathbb{K}[X]$. S désigne un sous-ensemble quelconque de E.

- $\mathcal{L}(E, F)$: espace vectoriel sur $\mathbb{K}$ des applications linéaires de E dans F.
- $\mathrm{Ker}\, f$: noyau de f (il s'agit d'un sous-espace vectoriel de E).
- $\mathrm{Im}\, f$: image de f (il s'agit d'un sous-espace vectoriel de F).
- $\dim(E)$: dimension de E sur $\mathbb{K}$ si E est de dimension finie.
- $\mathrm{Vect}\{v_1, \ldots, v_k\}$: sous-espace vectoriel de E engendré par $v_1, \ldots, v_k$.
- $\mathrm{Mat}_{\mathcal{B}}(v)$: vecteur colonne (matrice colonne) des coordonnées du vecteur v dans la base $\mathcal{B}$.
- ${}^tA$: transposée de la matrice A.
- Id_E : application identité de E.
- $\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(f)$: matrice représentative de f dans les bases $\mathcal{B}$ de E et $\mathcal{C}$ de F.
- $\mathrm{Mat}_{\mathcal{B}}(g)$: matrice représentative de l'endomorphisme g dans la base $\mathcal{B}$ de E.
- $P_{\mathcal{B}\to\mathcal{B}'}$: matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$ (il s'agit de la matrice $\mathrm{Mat}_{\mathcal{B}',\mathcal{B}}(\mathrm{Id}_E)$).
- $\mathrm{M}_{n,p}(\mathbb{K})$: espace vectoriel des matrices à n lignes et p colonnes et à coefficients dans $\mathbb{K}$.
- $\mathrm{M}_n(\mathbb{K})$: espace vectoriel des matrices carrées à n lignes et n colonnes et à coefficients dans $\mathbb{K}$.
- $\mathrm{GL}_n(\mathbb{K})$: groupe des matrices carrées inversibles de $\mathrm{M}_n(\mathbb{K})$.
- $\mathbb{K}_n[X]$: espace vectoriel sur $\mathbb{K}$ des polynômes en une indéterminée à coefficients dans $\mathbb{K}$ et de degré au plus n.

- $\deg(P)$: degré du polynôme P.
- $\mathrm{Tr}(M)$: trace de la matrice M i.e. la somme des coefficients diagonaux de M.
- I_n : matrice identité de taille n.
- 0_n : matrice nulle (i.e. avec uniquement des coefficients nuls) de $\mathrm{M}_n(\mathbb{K})$.
- $0_{n,p}$: matrice à n lignes et p colonnes avec uniquement des coefficients nuls.
- $\mathrm{Vect}(S)$: sous-espace vectoriel de E engendré par les vecteurs de E contenus dans S.
- $\mathcal{L}(E)$: espace vectoriel sur $\mathbb{K}$ des endomorphismes de E ($\mathcal{L}(E) = \mathcal{L}(E, E)$).
- $\mathrm{rg}(A)$: rang de la matrice A.
- $\mathrm{rg}\,(f)$: rang de l'application linéaire f.
- $\det(M)$: déterminant de la matrice M.
- $\det\,(g)$: déterminant de l'endomorphisme g.
- $E_{i\,j}$: matrice de $\mathrm{M}_n(\mathbb{K})$ avec un coefficient 1 sur la ligne i et la colonne j, et des coefficients nuls partout ailleurs.

Chapitre 1

Rappels et compléments sur la réduction des endomorphismes

1.1 Introduction

On donne dans ce chapitre des rappels et des compléments sur la théorie de la réduction des endomorphismes, c'est-à-dire l'étude des bases dans lesquelles un endomorphisme donné possède la représentation matricielle la plus "simple" possible (la plus "réduite" possible).

On étudiera notamment les critères nécessaires et suffisants classiques de diagonalisabilité, directs (via la recherche des espaces propres) ou utilisant la notion de polynôme d'endomorphisme. On étudiera également la trigonalisation et la réduction la plus aboutie des endomorphismes trigonalisables, à savoir la réduction de Jordan pour laquelle nous donnerons une méthode systématique de réduction. Nous énoncerons enfin le théorème de décomposition de Dunford d'un endomorphisme ou d'une matrice trigonalisable, utile notamment pour exprimer les puissances successives d'une matrice trigonalisable.

La première partie de ce chapitre étant constituée de rappels, les assertions seront la plupart du temps données sans preuve (on renvoie au cours de l'année passée pour les démonstrations). Nous les illustrerons cependant, ainsi que les méthodes, par des exemples. Dans la seconde partie du chapitre où l'on abordera des notions a priori nouvelles, quasiment toutes les preuves seront présentées ; seules resteront sous silence la preuve de l'unicité de la forme de Jordan d'un endomorphisme trigonalisable et la preuve de l'unicité de la décomposition de Dunford d'une matrice trigonalisable.

Tout au long de ce chapitre, $\mathbb{K}$ désigne un corps commutatif quelconque et E désigne un espace vectoriel sur $\mathbb{K}$ de dimension finie.

1.2 Valeurs propres et espaces propres

Soit f un endomorphisme de E.

Définition 1.2.1. *Soit $\lambda \in \mathbb{K}$. On dit que λ est une valeur propre de f s'il existe un vecteur non nul v de E tel que $f(v) = \lambda v \Leftrightarrow (f - \lambda \mathrm{Id}_E)(v) = 0_E$, autrement dit si l'endomorphisme*

$f - \lambda \mathrm{Id}_E$ n'est pas injectif.

Si $\lambda \in \mathbb{K}$ est une valeur propre de f, on note $E_\lambda := \mathrm{Ker}\ (f - \lambda \mathrm{Id}_E)$: il s'agit d'un sous-espace vectoriel de E que l'on appelle sous-espace propre de f associé à la valeur propre λ, et tout vecteur non nul de E_λ est appelé vecteur propre de f associé à la valeur propre λ. On dira qu'un vecteur v de E est un vecteur propre de f s'il existe une valeur propre $\lambda \in \mathbb{K}$ de f telle que v est un vecteur propre de f associé à λ.

Exemple 1.2.2. 2 est une valeur propre de l'endomorphisme

$$f : \begin{array}{ccc} \mathbb{R}^2 & \to & \mathbb{R}^2 \\ (x, y) & \mapsto & (x + 2y, -x + 4y) \end{array}$$

car, par exemple, $f(2, 1) = (4, 2) = 2(2, 1)$, et on a

$$E_2 = \mathrm{Ker}\ (f - 2\,\mathrm{Id}_{\mathbb{R}^2}) = \left\{(x, y) \in \mathbb{R}^2 \mid x = 2y\right\}.$$

Soit $n \in \mathbb{N}\backslash\{0\}$ et soit A une matrice carrée de $\mathrm{M}_n(\mathbb{K})$. On peut, de façon analogue, définir une notion de valeur propre, d'espace propre et de vecteur propre pour A : un scalaire $\lambda \in \mathbb{K}$ est une valeur propre de A s'il existe un vecteur colonne non nul V de $\mathrm{M}_{n,1}(\mathbb{K})$ tel que $AV = \lambda V$, autrement dit si le sous-espace vectoriel $E_\lambda := \mathrm{Ker}\ (A - \lambda I_n)$ de $\mathrm{M}_{n,1}(\mathbb{K})$ n'est pas réduit au vecteur colonne nul, et, dans ce cas, E_λ est appelé sous-espace propre de A associé à la valeur propre λ, et tout vecteur colonne non nul de E_λ est appelé vecteur propre de A associé à la valeur propre λ.

Remarque 1.2.3. Supposons que $\dim(E) = n$ et soit $\mathcal{B}$ une base de E. Soient $\lambda \in \mathbb{K}$ et $v \in E$. Si $A = \mathrm{Mat}_{\mathcal{B}}(f)$, alors λ est une valeur propre de f ssi λ est une valeur propre de A et, dans ce cas, v est un vecteur propre de f associé à λ ssi $\mathrm{Mat}_{\mathcal{B}}(v)$ est un vecteur propre de A associé à λ.

Définition 1.2.4. *L'ensemble des valeurs propres de f, resp. A, dans $\mathbb{K}$ est appelé spectre de f, resp. spectre de A, et noté* $\mathrm{Sp}(f)$*, resp.* $\mathrm{Sp}(A)$*.*

1.3 Polynôme caractéristique

Soient $f \in \mathcal{L}(E)$ et $\lambda \in \mathbb{K}$.

Proposition 1.3.1. *λ est une valeur propre de f ssi* $\det(f - \lambda \mathrm{Id}_E) = 0$.

On note χ_f le polynôme

$$\det(f - X\mathrm{Id}_E) \in \mathbb{K}[X],$$

appelé polynôme caractéristique de f. Ainsi, $\lambda \in \mathrm{Sp}(f)$ ssi λ est une racine (dans $\mathbb{K}$) de $\chi(f)$.

Exemple 1.3.2. Reprenons l'endomorphisme f de l'exemple 1.2.2. La matrice représentative de f dans la base canonique de $\mathbb{R}^2$ est

$$\begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix}$$

et donc

$$\chi_f = \det(f - X\mathrm{Id}_E) = \det\left(\begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix} - XI_2\right) = \det\begin{pmatrix} 1-X & 2 \\ -1 & 4-X \end{pmatrix} = (3-X)(2-X)$$

d'où $\mathrm{Sp}(f) = \{2; 3\}$.

Soit $A \in \mathrm{M}_n(\mathbb{K})$, on définit de façon analogue le polynôme caractéristique $\chi_A := \det(A - XI_n) \in \mathbb{K}[X]$ de A, et le spectre de A est alors l'ensemble des racines de χ_A dans $\mathbb{K}$.

Remarque 1.3.3. Attention au corps de base : si l'on considère par exemple la matrice

$$A := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

on a

$$\chi_A = \det\begin{pmatrix} -X & 1 \\ -1 & -X \end{pmatrix} = X^2 + 1.$$

Ainsi, si A est considérée comme une matrice de $\mathrm{M}_n(\mathbb{C})$, on a $\mathrm{Sp}(A) = \mathrm{Sp}_{\mathbb{C}}(A) = \{-i; i\}$, et si A est considérée comme une matrice de $\mathrm{M}_n(\mathbb{R})$, on a $\mathrm{Sp}(A) = \mathrm{Sp}_{\mathbb{R}}(A) = \varnothing$.

Proposition 1.3.4. *On a*

$$\chi_A = (-1)^n X^n + (-1)^{n-1}\mathrm{Tr}(A)X^{n-1} + \cdots + \det(A).$$

En particulier, le polynôme caractéristique de A est de degré n et A possède donc au plus n valeurs propres distinctes.

Corollaire 1.3.5. *Si $n = \dim(E)$, le polynôme $\chi_f \in \mathbb{K}[X]$ est de degré n et f possède au plus n valeurs propres distinctes.*

Remarque 1.3.6. D'après le théorème de d'Alembert-Gauss, tout polynôme de $\mathbb{C}[X]$ est scindé. Ainsi, tout endomorphisme sur $\mathbb{C}$, resp. toute matrice de $\mathrm{M}_n(\mathbb{C})$, admet au moins une valeur propre.

1.4 Diagonalisabilité et diagonalisation

Soit $f \in \mathcal{L}(E)$. On note $n := \dim(E)$.

La forme la plus "simple" que peut avoir une matrice carrée est la forme diagonale. Nous allons rappeler dans cette section des conditions nécessaires et suffisantes sous lesquelles f possède une matrice représentative diagonale (diagonalisabilité). Dans ce cas, on cherchera à déterminer une base de E dans laquelle la matrice représentative de f est diagonale (diagonalisation).

On commence par donner la définition précise de la diagonalisabilité de f :

Définition 1.4.1. *On dit que l'endomorphisme f est diagonalisable s'il existe une base $\mathcal{B}$ de E et des scalaires $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ tels que*

$$\mathrm{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

Remarquons que f est donc diagonalisable si et seulement s'il existe une base de E formée de vecteurs propres de f. Mais on peut énoncer une caractérisation plus utile en pratique. Pour cela, commençons par énoncer le fait suivant :

Proposition 1.4.2. *Soient $\lambda_1, \dots, \lambda_k$, $k \in \mathbb{N}\backslash\{0\}$, des valeurs propres deux à deux distinctes de f. Alors les sous-espaces propres $E_{\lambda_1}, \dots, E_{\lambda_k}$ correspondants sont en somme directe.*

En conséquence, si $\lambda_1, \dots, \lambda_p$, $p \in \{0, \dots, n\}$, désignent les valeurs propres deux à deux distinctes de f :

Théorème 1.4.3. *f est diagonalisable ssi $E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_n} = E$ ssi $\displaystyle\sum_{i=1}^{p} \dim(E_{\lambda_i}) = \dim(E)$.*

Exemple 1.4.4. Reprenons l'endomorphisme f des exemples 1.2.2 et 1.3.2. On avait déterminé que $\mathrm{Sp}(f) = \{2; 3\}$ et $E_2 = \left\{(x, y) \in \mathbb{R}^2 \mid x = 2y\right\} = \mathrm{Vect}\{(2, 1)\}$. Enfin,

$$E_3 = \mathrm{Ker}\ (f - 3\,\mathrm{Id}_{\mathbb{R}^2}) = \left\{(x, y) \in \mathbb{R}^2 \mid x = y\right\} = \mathrm{Vect}\{(1, 1)\}.$$

Ainsi, $\dim(E_1) + \dim(E_2) = 2 = \dim\left(\mathbb{R}^2\right)$ et f est donc diagonalisable. De plus, la famille $\mathcal{B} := \{(2, 1), (1, 1)\}$ est une base de E formée de vecteurs propres de f et on a

$$\mathrm{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$$

(on dit qu'on a diagonalisé f).

Remarque 1.4.5. Diagonaliser un endomorphisme diagonalisable f de E, c'est déterminer une base $\mathcal{B}$ de E dans laquelle la matrice représentative de f est diagonale et exprimer $\mathrm{Mat}_{\mathcal{B}}(f)$.

On peut également exprimer la condition sur les dimensions du théorème 1.4.3 à l'aide du polynôme caractéristique et plus particulièrement à l'aide des multiplicités des valeurs propres en tant que racines du polynôme caractéristique :

Définition 1.4.6. *Soit $\lambda \in \mathbb{K}$ une valeur propre de f. On note m_λ la multiplicité de λ en tant que racine du polynôme $\chi_f \in \mathbb{K}[X]$.*

On peut tout de suite remarquer que, si $\lambda_1, \dots, \lambda_p$ désignent les valeurs propres de f, alors la somme $m_{\lambda_1} + \cdots + m_{\lambda_p}$ est inférieure ou égale à $\deg(\chi_f) = n = \dim(E)$, et qu'il y a égalité ssi χ_f est scindé sur $\mathbb{K}$. De plus :

Proposition 1.4.7. *Si $\lambda \in \mathrm{Sp}(f)$, alors*

$$1 \leqslant \dim(E_\lambda) \leqslant m_\lambda.$$

Ainsi :

Théorème 1.4.8. *f est diagonalisable ssi χ_f est scindé et, pour tout $\lambda \in \mathrm{Sp}(f)$, $\dim(E_\lambda) = m_\lambda$.*

Exemple 1.4.9. Si f admet $n = \dim(E)$ valeurs propres deux à deux distinctes, alors f est diagonalisable.

La diagonalisation d'un endomorphisme correspond à un changement de base vers une base dans laquelle la matrice représentative de l'endomorphisme considéré est diagonale. L'analogue matriciel du changement de base est l'opération de "conjugaison" par une matrice inversible. Soit A une matrice de $\mathrm{M}_n(\mathbb{K})$.

Définition 1.4.10. *On dit que A est diagonalisable s'il existe une matrice inversible $P \in \mathrm{GL}_n(\mathbb{K})$ et une matrice diagonale D de $\mathrm{M}_n(\mathbb{K})$ telles que*

$$P^{-1}AP = D.$$

Remarque 1.4.11. Si $B \in \mathrm{M}_n(\mathbb{K})$, on dit que A est semblable à B s'il existe une matrice inversible $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $P^{-1}AP = B$ (la relation de similitude sur $\mathrm{M}_n(\mathbb{K})$ est une relation d'équivalence).

Ainsi, A est diagonalisable ssi A est semblable à une matrice diagonale.

Les résultats de diagonalisabilité d'un endomorphisme énoncés ci-dessus ont leurs analogues matriciels, à savoir :

Théorème 1.4.12. *Soient $\lambda_1, \ldots, \lambda_p$, $p \in \{0, \ldots, n\}$, les valeurs propres deux à deux distinctes de A et, pour $i \in \{1, \ldots, p\}$, notons m_{λ_i} la multiplicité de λ_i en tant que racine de χ_A. Alors A est diagonalisable ssi $\sum_{i=1}^{p} \dim(E_{\lambda_i}) = n$ ssi (χ_A est scindé et, pour tout $i \in \{1, \ldots, p\}$, $\dim(E_{\lambda_i}) = m_{\lambda_i}$).*

Remarque 1.4.13.

- Diagonaliser une matrice diagonalisable A de $\mathrm{M}_n(\mathbb{K})$, c'est déterminer une matrice inversible $P \in \mathrm{GL}_n(\mathbb{K})$ telle que la matrice $P^{-1}AP$ soit diagonale et exprimer $P^{-1}AP$.
- Diagonaliser une matrice diagonalisable permet entre autres choses de calculer ses puissances, comme présenté dans l'exemple ci-dessous.

Exemple 1.4.14. On considère la matrice

$$A := \begin{pmatrix} 1 & -1 \\ 2 & 4 \end{pmatrix}$$

de $\mathrm{M}_2(\mathbb{R})$. Son polynôme caractéristique $\chi_A = (X-2)(X-3)$ est scindé à racines simples donc A est diagonalisable.

Une base de E_2 est $\left\{ \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$, une base de E_3 est $\left\{ \begin{pmatrix} 1 \\ -2 \end{pmatrix} \right\}$ et on pose

$$P := \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix}.$$

On a alors

$$\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} = P^{-1}AP$$

donc $A = P \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} P^{-1}$ et, par associativité du produit matriciel, pour $k \in \mathbb{N} \backslash \{0\}$, $A^k = P \begin{pmatrix} 2^k & 0 \\ 0 & 3^k \end{pmatrix} P^{-1}$.

Or $P^{-1} = \begin{pmatrix} 2 & 1 \\ -1 & -1 \end{pmatrix}$ donc

$$A^k = \begin{pmatrix} 2^{k+1} - 3^k & 2^k + 3^k \\ -2^{k+1} + 2 \cdot 3^k & -2^k - 2 \cdot 3^k \end{pmatrix}.$$

1.5 Trigonalisabilité et trigonalisation

Soit $f \in \mathcal{L}(E)$.

Si f n'est pas diagonalisable, on peut chercher à déterminer si f peut être réduit sous forme triangulaire :

Définition 1.5.1. *On dit que l'endomorphisme f est trigonalisable (on dit aussi triangularisable) s'il existe une base $\mathcal{B}$ de E telle que $\mathrm{Mat}_{\mathcal{B}}(f)$ est triangulaire (supérieure ou inférieure).*

De manière analogue, on dira qu'une matrice A de $\mathrm{M}_n(\mathbb{K})$ est trigonalisable si A est semblable à une matrice triangulaire.

Remarque 1.5.2.
- Si T est une matrice triangulaire représentant f, les coefficients de la diagonale de T sont les valeurs propres de f, apparaissant suivant leurs multiplicités dans χ_f.
- Un endomorphisme diagonalisable est en particulier trigonalisable.

On donne dès à présent une condition nécessaire et suffisante de trigonalisabilité d'un endomorphisme :

Théorème 1.5.3. *L'endomorphisme f est trigonalisable si et seulement si son polynôme caractéristique χ_f est scindé sur $\mathbb{K}$.*

Corollaire 1.5.4. *Tout endomorphisme d'un $\mathbb{C}$-espace vectoriel est trigonalisable.*

Toute matrice de $\mathrm{M}_n(\mathbb{C})$ est semblable à une matrice triangulaire.

Exemple 1.5.5. Considérons la matrice

$$A = \begin{pmatrix} 3 & -1 & 1 \\ 2 & 0 & 1 \\ 1 & -1 & 2 \end{pmatrix}$$

de $\mathrm{M}_3(\mathbb{R})$. Son polynôme caractéristique est $\chi_A = -(X-1)(X-2)^2$ et est donc scindé sur $\mathbb{R}$. La matrice A est donc trigonalisable d'après le théorème ci-dessus. Cherchons une matrice triangulaire T de $\mathrm{M}_2(\mathbb{R})$ semblable à A.

Commençons par déterminer les espaces propres de A. On a $A - I_3 = \begin{pmatrix} 2 & -1 & 1 \\ 2 & -1 & 1 \\ 1 & -1 & 1 \end{pmatrix}$ donc

$$E_1 = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathrm{M}_3(\mathbb{R}) \mid \begin{cases} 2x - y + z & = 0 \\ x - y + z & = 0 \end{cases} \right\} = \mathrm{Vect}\left\{ \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\}$$

et $A - 2I_3 = \begin{pmatrix} 1 & -1 & 1 \\ 2 & -2 & 1 \\ 1 & -1 & 0 \end{pmatrix}$ donc

$$E_2 = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathrm{M}_3(\mathbb{R}) \mid \begin{cases} x - y + z & = 0 \\ x - y & = 0 \end{cases} \right\} = \mathrm{Vect}\left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\}$$

(remarquons ici que $\dim(E_2) < m_2$ donc A n'est pas diagonalisable).

Si on complétait maintenant la famille libre $\left\{ \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\}$ de $\mathrm{M}_{3,1}(\mathbb{R})$ ainsi obtenue en une base, i.e. de façon à ce que la matrice

$$P := \begin{pmatrix} 0 & 1 & \star \\ 1 & 1 & \star \\ 1 & 0 & \star \end{pmatrix}$$

soit inversible, on aurait alors

$$P^{-1}AP = \begin{pmatrix} 1 & 0 & \star \\ 0 & 2 & \star \\ 0 & 0 & 2 \end{pmatrix}$$

(deux matrices semblables ont les mêmes polynômes caractéristiques).

Ainsi, dans ce cas, si on note $V_1 := \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ et $V_2 := \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, n'importe quel vecteur colonne V_3 de $\mathrm{M}_{3,1}(\mathbb{R})$ n'appartenant pas à $\mathrm{Vect}\{V_1, V_2\}$ convient. On choisit, <u>par exemple</u>, $V_3 := \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, on pose alors

$$P := \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$

qui est bien une matrice inversible, et on a

$$P^{-1}AP = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

Remarque 1.5.6. Si on avait posé $V_3 := \begin{pmatrix}1\\1\\1\end{pmatrix}$ et $P := \begin{pmatrix}0&1&1\\1&1&1\\1&0&1\end{pmatrix}$, on aurait eu $P^{-1}AP = \begin{pmatrix}1&0&0\\0&2&1\\0&0&2\end{pmatrix}$. En particulier, on peut trigonaliser A de plusieurs façons.

Remarque 1.5.7.

1. Trigonaliser un endomorphisme trigonalisable f de E, c'est déterminer une base $\mathcal{B}$ de E dans laquelle la matrice représentative de f est triangulaire et exprimer $\mathrm{Mat}_{\mathcal{B}}(f)$.
2. Trigonaliser une matrice trigonalisable A de $\mathrm{M}_n(\mathbb{K})$, c'est déterminer une matrice inversible $P \in \mathrm{GL}_n(\mathbb{K})$ telle que la matrice $P^{-1}AP$ soit triangulaire et exprimer $P^{-1}AP$.

Si l'endomorphisme f est trigonalisable, il existe une trigonalisation plus “simple” que les autres : la réduction sous forme de Jordan. Celle-ci fera l'objet d'une section ultérieure de ce chapitre et donnera lieu à une procédure algorithmique de trigonalisation. Cette méthode repose notamment sur la notion de polynôme d'endomorphisme que nous allons introduire dans la section suivante.

1.6 Polynômes d'endomorphismes et polynômes annulateurs

Nous allons pouvoir exprimer de nouvelles conditions nécessaires et suffisantes de réductibilité des endomorphismes grâce à la notion de polynôme d'endomorphisme :

Définition 1.6.1. *Soient $P = a_N X^N + a_{N-1}X^{N-1} + \cdots + a_1 X + a_0 \in \mathbb{K}[X]$ et $f \in \mathcal{L}(E)$. On note*

$$P(f) := a_N f^N + a_{N-1} f^{N-1} + \cdots + a_1 f + a_0 \mathrm{Id}_E \in \mathcal{L}(E)$$

(où, pour $k \in \{0, \ldots, N\}$, f^k désigne la composée $k^{\text{ème}}$ de l'endomorphisme f avec lui-même). Un endomorphisme de E de cette forme est appelé polynôme d'endomorphisme.

Exemple 1.6.2. Si $f \in \mathcal{L}(E)$ et si P est le polynôme $3X+2$ de $\mathbb{R}[X]$, on a $P(f) = (3X+2)(f) = 3f + 2\mathrm{Id}_E$.

On peut définir de façon analogue la notion de polynôme de matrice : si l'on reprend les notations de la définition ci-dessus et si A est une matrice de $\mathrm{M}_n(\mathbb{K})$, on définit

$$P(A) := a_N A^N + a_{N-1}A^{N-1} + \cdots + a_1 A + a_0 I_n \in \mathrm{M}_n(\mathbb{K}),$$

où, pour $k \in \{0, \ldots, N\}$, A^k désigne la puissance $k^{\text{ème}}$ de A.

Exemple 1.6.3. Considérons la matrice

$$A := \begin{pmatrix}1&2\\-1&4\end{pmatrix}$$

de $\mathrm{M}_2(\mathbb{R})$ et le polynôme $P := -X^2 + X - 4$ de $\mathbb{R}[X]$. On a

$$P(A) = \left(-X^2 + X - 4\right)(A) = -A^2 + A - 4I_2 = -\begin{pmatrix}-1&10\\-5&14\end{pmatrix} + \begin{pmatrix}1&2\\-1&4\end{pmatrix} - 4\begin{pmatrix}1&0\\0&1\end{pmatrix} = \begin{pmatrix}-2&-8\\4&-14\end{pmatrix}.$$

Remarque 1.6.4. Si E est de dimension finie n, si $\mathcal{B}$ est une base de E et si $f \in \mathcal{L}(E)$, alors, pour tout polynôme $P \in \mathbb{K}[X]$,

$$\mathrm{Mat}_{\mathcal{B}}\left(P(f)\right) = P\left(\mathrm{Mat}_{\mathcal{B}}(f)\right).$$

Nous allons à présent établir les propriétés de base des polynômes d'endomorphismes. Tous les énoncés et notions présentés ci-après sur les polynômes d'endomorphismes ont leurs analogues immédiats pour les polynômes de matrices.

Soit $f \in \mathcal{L}(E)$.

Proposition 1.6.5. *Pour tous polynômes $P, Q \in \mathbb{K}[X]$ et tous scalaires $\lambda, \mu \in \mathbb{K}$, on a*

$$(\lambda P + \mu Q)(f) = \lambda P(f) + \mu Q(f) \quad \text{et} \quad (PQ)(f) = P(f) \circ Q(f).$$

Démonstration. Notons $N := \max\left(\deg(P), \deg(Q)\right)$ et écrivons $P = \sum_{i=0}^{N} a_i X^i$ et $Q = \sum_{j=0}^{N} b_j X^j$. On a

$$\begin{aligned}
(\lambda P + \mu Q)(f) &= \left(\sum_{i=0}^{N} (\lambda a_i + \mu b_i) X^i\right)(f) \\
&= \sum_{i=0}^{N} (\lambda a_i + \mu b_i) f^i \\
&= \lambda \left(\sum_{i=0}^{N} a_i f^i\right) + \mu \left(\sum_{j=0}^{N} b_j f^j\right) \\
&= \lambda P(f) + \mu Q(f)
\end{aligned}$$

et

$$\begin{aligned}
P(f) \circ Q(f) &= \left(\sum_{i=0}^{N} a_i f^i\right) \circ \left(\sum_{j=0}^{N} b_j f^j\right) \\
&= \sum_{i=0}^{N} \sum_{j=0}^{N} a_i b_j f^i \circ f^j \ (f \text{ est une application linéaire}) \\
&= \sum_{0 \leqslant i,j \leqslant N} a_i b_j f^{i+j} \\
&= \sum_{k=0}^{2N} \left(\sum_{i+j=k} a_i b_j f^k\right) \\
&= \sum_{k=0}^{2N} \left(\sum_{i+j=k} a_i b_j\right) f^k \\
&= \left(\sum_{k=0}^{2N} \left(\sum_{i+j=k} a_i b_j\right) X^k\right)(f) \\
&= (PQ)(f).
\end{aligned}$$

□

Exemple 1.6.6. Considérons la matrice

$$A := \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix}$$

de $\mathrm{M}_2(\mathbb{R})$ et le polynôme $P := X(X-1)$ de $\mathbb{R}[X]$. On a

$$P(A) = A(A - I_2) = \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ -1 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Remarque 1.6.7.

- Si $P, Q \in \mathbb{K}[X]$, $P(f) \circ Q(f) = (PQ)(f) = (QP)(f) = Q(f) \circ P(f)$.
- Une autre façon d'exprimer la proposition 1.6.5 est de dire que l'application

$$\begin{array}{ccc} \mathbb{K}[X] & \to & \mathcal{L}(E) \\ P & \mapsto & P(f) \end{array}$$

est un morphisme de $\mathbb{K}$-algèbres.

Introduisons ensuite la notion de polynôme annulateur d'un endomorphisme. Soit $P \in \mathbb{K}[X]$.

Définition 1.6.8. *On dit que P est un polynôme annulateur de f (ou que le polynôme P annule l'endomorphisme f) si $P(f)$ est l'endomorphisme identiquement nul $0_{\mathcal{L}(E)}$ de E.*

Exemple 1.6.9.

- Si l'on reprend les notations de l'exemple précédent 1.6.6, le polynôme $P = X(X-1)$ est un polynôme annulateur de la matrice $A = \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix}$.
- Le polynôme $X^2 - X$ annule f si et seulement si f est une projection. En effet,

$$\begin{array}{lll} f \text{ est une projection de } E & \text{ssi} & f^2 = f \\ & \text{ssi} & f^2 - f = 0_{\mathcal{L}(E)} \\ & \text{ssi} & (X^2 - X)(f) = 0_{\mathcal{L}(E)} \\ & \text{ssi} & X^2 - X \text{ est un polynôme annulateur de } f. \end{array}$$

- Le polynôme $X^2 - 1$ annule f si et seulement si f est une symétrie. En effet,

$$\begin{array}{lll} f \text{ est une symétrie de } E & \text{ssi} & f^2 = \mathrm{Id}_E \\ & \text{ssi} & f^2 - \mathrm{Id}_E = 0_{\mathcal{L}(E)} \\ & \text{ssi} & (X^2 - 1)(f) = 0_{\mathcal{L}(E)} \\ & \text{ssi} & X^2 - 1 \text{ est un polynôme annulateur de } f. \end{array}$$

Remarque 1.6.10. Comme E est de dimension finie, tout endomorphisme de E admet un polynôme annulateur non nul. En effet, si n désigne la dimension de E, le $\mathbb{K}$-espace vectoriel $\mathcal{L}(E)$ est de dimension finie égale à n^2 et la famille $\left\{\mathrm{Id}_E, f, f^2, \ldots, f^{n^2}\right\}$ de vecteurs de $\mathcal{L}(E)$ de cardinal n^2+1 est liée : il existe $(a_0,\ldots,a_{n^2}) \in \mathbb{K}^{n^2+1}\backslash\{(0,\ldots,0)\}$ tel que $a_0\mathrm{Id}_E + a_1 f + \ldots + a_{n^2} f^{n^2} = 0_{\mathcal{L}(E)}$ et le polynôme non nul $\sum_{i=0}^{n^2} a_i X^i$ annule donc l'endomorphisme f.

Un premier lien entre les polynômes annulateurs et la réduction des endomorphismes est donné par la proposition suivante et son corollaire :

Proposition 1.6.11. *Soit $\lambda \in \mathrm{Sp}(f)$ une valeur propre de f et soit $v \in E_\lambda$. Alors, pour tout $P \in \mathbb{K}[X]$, $\big(P(f)\big)(v) = P(\lambda)\cdot v$ (où $\cdot$ désigne ici le produit d'un vecteur de E par un scalaire de $\mathbb{K}$).*

Démonstration. On commence par montrer par récurrence que, pour tout $k \in \mathbb{N}$, $f^k(v) = \lambda^k v$. La propriété est vraie au rang $k=0$ car $f^0(v) = \mathrm{Id}_E(v) = v = \lambda^0 v$, et, si l'on suppose la propriété vraie au rang k pour un entier $k \in \mathbb{N}$ fixé, on a

$$\begin{aligned}
f^{k+1}(v) &= f\left(f^k(v)\right) \\
&= f\left(\lambda^k v\right) \text{ (par hypothèse de récurrence)} \\
&= \lambda^k f(v) \text{ (car } f \text{ est une application linéaire)} \\
&= \lambda^k \cdot (\lambda v) \text{ (car } v \in E_\lambda) \\
&= \lambda^{k+1} v.
\end{aligned}$$

Soit maintenant $P = \sum_{k=0}^{N} a_k X^k$ un polynôme de $\mathbb{K}[X]$, on a

$$\begin{aligned}
\big(P(f)\big)(v) &= \left(\sum_{k=0}^{N} a_k f^k\right)(v) \\
&= \sum_{k=0}^{N} a_k f^k(v) \\
&= \sum_{k=0}^{N} a_k \cdot \left(\lambda^k v\right) \\
&= \sum_{k=0}^{N} \left(a_k \lambda^k\right) \cdot v \\
&= \left(\sum_{k=0}^{N} a_k \lambda^k\right) \cdot v \\
&= P(\lambda)\cdot v.
\end{aligned}$$

□

Corollaire 1.6.12. *Soit $\lambda \in \mathrm{Sp}\,(f)$. Si P est un polynôme annulateur de f, alors λ est une racine de P.*

Démonstration. Soit v un vecteur propre de f pour la valeur propre λ (en particulier, v est différent du vecteur nul 0_E de E). On a alors, par la proposition précédente 1.6.11,

$$\begin{aligned} P(\lambda) \cdot v &= \big(P(f)\big)(v) \\ &= 0_{\mathcal{L}(E)}(v) \text{ (P est un polynôme annulateur de f)} \\ &= 0_E. \end{aligned}$$

Or $v \neq 0_E$ donc, nécessairement, $P(\lambda) = 0$. □

Exemple 1.6.13.
- Reprenons la matrice A de l'exemple 1.6.6. Comme $P = X(X-1)$ est un polynôme annulateur de A, $\mathrm{Sp}(A) \subset \{0; 1\}$.
- Si l'endomorphisme f vérifie $f^3 = f$ i.e. le polynôme $X^3 - X = X(X-1)(X+1)$ annule f, alors $\mathrm{Sp}(f) \subset \{0; -1; 1\}$.

Remarque 1.6.14. Toute racine d'un polynôme annulateur de f n'est pas nécessairement une valeur propre de f : par exemple, le polynôme $X(X-1)$ annule Id_E bien que 0 ne soit pas une valeur propre de Id_E.

On en vient à un premier critère, nécessaire et suffisant, de diagonalisabilité d'un endomorphisme mettant en jeu la notion de polynôme annulateur :

Théorème 1.6.15. *L'endomorphisme f est diagonalisable si et seulement s'il existe un polynôme annulateur de f dans $\mathbb{K}[X]$ qui soit scindé à racines simples.*

Démonstration. Notons tout d'abord $\lambda_1, \dots, \lambda_p$, $p \in \{0, \dots, n\}$, les valeurs propres deux à deux distinctes de f et, pour tout $i \in \{1, \dots, p\}$, $d_i := \dim(E_{\lambda_i})$.

Commençons par montrer l'implication directe : on suppose que f est diagonalisable. Il existe donc une base $\mathcal{B}$ de E dans laquelle la matrice représentative de f est la matrice

$$\mathrm{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} \lambda_1 I_{d_1} & & 0 \\ & \ddots & \\ 0 & & \lambda_p I_{d_p} \end{pmatrix}.$$

Notons A cette matrice. Si l'on pose $P := \displaystyle\prod_{i=1}^{p} (X - \lambda_i) \in \mathbb{K}[X]$, P est un polynôme scindé à racines simples de $\mathbb{K}[X]$ et

$$\begin{aligned} P(A) &= \prod_{i=1}^{p} (A - \lambda_i I_n) \\ &= \begin{pmatrix} \boxed{0_{d_1}} & & & 0 \\ & \boxed{(\lambda_2 - \lambda_1) I_{d_2}} & & \\ & & \ddots & \\ 0 & & & \boxed{(\lambda_p - \lambda_1) I_{d_p}} \end{pmatrix} \cdots \begin{pmatrix} \boxed{(\lambda_1 - \lambda_p) I_{d_1}} & & & 0 \\ & \ddots & & \\ & & \boxed{(\lambda_{p-1} - \lambda_p) I_{d_p}} & \\ 0 & & & \boxed{0_{d_p}} \end{pmatrix} \\ &= 0_n. \end{aligned}$$

Réciproquement, supposons qu'il existe un polynôme scindé à racines simples $P = \alpha \prod_{j=1}^{m} (X - \mu_j) \in \mathbb{K}[X]$, avec $\alpha \in \mathbb{K}\backslash\{0\}$ et $\mu_1, \ldots, \mu_m \in \mathbb{K}$ deux à deux distincts, qui soit annulateur de f.

D'après le corollaire 1.6.12, $\mathrm{Sp}(f) \subset \{\mu_1, \ldots, \mu_m\}$ et, pour tout $j \in \{1, \ldots, m\}$,

$$\mathrm{Ker}\ (f - \mu_j \mathrm{Id}_E) = \begin{cases} E_{\mu_j} & \text{si } \mu_j \in \mathrm{Sp}(f), \\ \{0_E\} & \text{sinon.} \end{cases}.$$

Ainsi,

$$\sum_{j=1}^{m} \mathrm{Ker}\ (f - \mu_j \mathrm{Id}_E) = \sum_{i=1}^{p} E_{\lambda_i} = \bigoplus_{i=1}^{p} E_{\lambda_i}.$$

Or f est diagonalisable si et seulement si $\bigoplus_{i=1}^{p} E_{\lambda_i} = E$: montrons donc que $\sum_{j=1}^{m} \mathrm{Ker}\ (f - \mu_j \mathrm{Id}_E) = E$.

Pour tout $k \in \{1, \ldots, m\}$, on commence par noter

$$P_k := \prod_{j \neq k} (X - \mu_j) \text{ et } \alpha_k := \frac{1}{P_k(\mu_k)}$$

($P_k(\mu_k) = \prod_{j \neq k} (\mu_k - \mu_j) \neq 0$ car les scalaires $\mu_1, \ldots, \mu_m$ sont deux à deux distincts).

On a alors $\sum_{k=1}^{m} \alpha_k P_k = 1$ dans $\mathbb{K}[X]$. En effet le polynôme $1 - \sum_{k=1}^{m} \alpha_k P_k$ est de degré au plus $m-1$ (puisque les polynômes $P_1, \ldots, P_m$ le sont) et possède m racines (pour tout $j \in \{1, \ldots, m\}$, $1 - \sum_{k=1}^{m} \alpha_k P_k(\mu_j) = 1 - \alpha_j P_j(\mu_j) = 0$) donc est le polynôme nul.

Ainsi,

$$\sum_{k=1}^{m} \alpha_k P_k(f) = 1(f) = \mathrm{Id}_E.$$

Soit maintenant $v \in E$. On a $v = \mathrm{Id}_E(v) = \sum_{k=1}^{m} \alpha_k P_k(f)(v)$. Pour tout $k \in \{1, \ldots, m\}$, on note $v_k := \alpha_k P_k(f)(v)$, de sorte que $v = \sum_{k=1}^{m} v_k$. Fixons alors $k \in \{1, \ldots, m\}$ et montrons que $v_k \in \mathrm{Ker}\ (f - \mu_k \mathrm{Id}_E)$, ce qui prouvera que $E = \sum_{k=1}^{m} \mathrm{Ker}\ (f - \mu_k \mathrm{Id}_E)$ et conclura ainsi la démonstration.

On a

$$\begin{aligned}
(f - \mu_k \mathrm{Id}_E)(v_k) &= \Big(f - \mu_k \mathrm{Id}_E\Big)\big(\alpha_k P_k(f)(v)\big) \\
&= \big(X - \mu_k\big)(f) \circ \big(\alpha_k P_k\big)(f)(v) \\
&= \Big((X - \mu_k)(\alpha_k P_k)\Big)(f)(v) \\
&= \big(\alpha_k \alpha^{-1} P\big)(f)(v) \\
&= 0_{\mathcal{L}(E)}(v) \text{ (car } P \text{ annule l'endomorphisme } f) \\
&= 0_E,
\end{aligned}$$

et donc $v_k \in \mathrm{Ker}\ (f - \mu_k \mathrm{Id}_E)$. □

Ce critère permet, si l'on trouve un tel polynôme annulateur de f scindé à racines simples, de montrer que f est diagonalisable sans avoir à calculer les dimensions des espaces propres de f :

Exemple 1.6.16. Si f vérifie alors $f^3 = f$ alors f est diagonalisable car le polynôme $X^3 - X = X(X-1)(X+1)$, annulateur de f, est scindé à racines simples.

Remarque 1.6.17. D'après la preuve de l'équivalence du théorème 1.6.15, f est diagonalisable ssi le polynôme $\prod_{\lambda \in \mathrm{Sp}(f)} (X - \lambda) \in \mathbb{K}[X]$ (qui est scindé à racines simples) annule f.

1.7 Polynôme minimal

Soit $f \in \mathcal{L}(E)$. Remarquons que l'ensemble I_f des polynômes de $\mathbb{K}[X]$ qui annulent l'endomorphisme f est un idéal de l'anneau $\mathbb{K}[X]$: il s'agit du noyau du morphisme d'anneaux $\mathbb{K}[X] \to \mathcal{L}(E)$; $P \mapsto P(f)$ (cf. remarque 1.6.7).

Or tout idéal de l'anneau principal $\mathbb{K}[X]$ peut être engendré par un unique élément unitaire :

Proposition et Définition 1.7.1. *Il existe un unique polynôme unitaire μ_f, annulateur de f, tel que $I_f = (\mu_f) = \{\mu_f Q \mid Q \in \mathbb{K}[X]\}$. On appelle μ_f le polynôme minimal de f.*

Remarque 1.7.2.

- Les polynômes annulateurs de f sont exactement les multiples de μ_f. En particulier, si l'on connaît μ_f, on connaît tous les polynômes annulateurs de f.

- μ_f est le polynôme annulateur de f unitaire non nul de plus petit degré (comme E est de dimension finie, f possède un polynôme annulateur non nul par la remarque 1.6.10). En particulier, $\deg(\mu_f) \geqslant 1$.

- Si A est une matrice de $\mathrm{M}_n(\mathbb{K})$, $n \in \mathbb{N}\backslash\{0\}$, on peut définir de façon analogue le polynôme minimal μ_A de A. Toutes les propriétés sur le polynôme minimal d'un endomorphisme ont leurs analogues pour le polynôme minimal d'une matrice carrée.

- Si $n := \dim(E)$, si $\mathcal{B}$ est une base de E et si A désigne la matrice représentative $\mathrm{Mat}_{\mathcal{B}}(f)$ de f dans la base $\mathcal{B}$, $\mu_f = \mu_A$. En effet,
$$\begin{aligned} \mu_f(A) &= \mu_f\left(\mathrm{Mat}_{\mathcal{B}}(f)\right) \\ &= \mathrm{Mat}_{\mathcal{B}}\left(\mu_f(f)\right) \text{ (remarque 1.6.4)} \\ &= \mathrm{Mat}_{\mathcal{B}}\left(0_{\mathcal{L}(E)}\right) \\ &= 0_n \end{aligned}$$
donc μ_A divise μ_f. Réciproquement,
$$\begin{aligned} \mathrm{Mat}_{\mathcal{B}}\left(\mu_A(f)\right) &= \mu_A\left(\mathrm{Mat}_{\mathcal{B}}(f)\right) \text{ (remarque 1.6.4)} \\ &= \mu_A(A) \\ &= 0_n \end{aligned}$$
donc $\mu_A(f) = 0_{\mathcal{L}(E)}$ donc μ_f divise μ_A. Comme les polynômes μ_f et μ_A sont de plus tous deux unitaires, on en déduit que $\mu_f = \mu_A$.

Le polynôme minimal de f divise le polynôme caractéristique de f en vertu du théorème de Cayley-Hamilton :

Théorème 1.7.3 (Théorème de Cayley-Hamilton)**.** *Le polynôme caractéristique de f est un polynôme annulateur de f. Autrement dit $\chi_f \in I_f$ i.e. μ_f divise χ_f.*

Pour montrer le théorème de Cayley-Hamilton, nous utiliserons les deux résultats suivants :

Lemme 1.7.4. *Soit F un sous-espace vectoriel de E tel que $f(F) \subset F$ (on dit que F est stable par f). Si l'on note $g : F \to F$; $v \mapsto f(v)$ la restriction de f à F, alors le polynôme caractéristique χ_g de l'endomorphisme g de F divise le polynôme caractéristique χ_f de l'endomorphisme f de E.*

Démonstration. Soit $\mathcal{B}_0$ une base de F. On complète la famille libre $\mathcal{B}_0$ de F en une base $\mathcal{B}$ de E. La matrice représentative de $f \in \mathcal{L}(E)$ dans la base $\mathcal{B}$ est alors
$$A := \mathrm{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} \mathrm{Mat}_{\mathcal{B}_0}(g) & \star \\ 0 & B \end{pmatrix}$$
où B est une matrice carrée de taille $n-p$ si $p := \dim(F)$ (on utilise ici le fait que F est stable par f : l'image par f de tout vecteur de $\mathcal{B}_0$ est dans F donc est une combinaison linéaire de vecteurs de $\mathcal{B}_0$).

Ainsi, si l'on note $A_0 := \mathrm{Mat}_{\mathcal{B}_0}(g) \in \mathrm{M}_p(\mathbb{K})$,
$$\begin{aligned} \chi_f &= \chi_A \\ &= \det\left(A - XI_n\right) \\ &= \det\begin{pmatrix} A_0 - XI_p & \star \\ 0 & B - XI_{n-p} \end{pmatrix} \\ &= \det\left(A_0 - XI_p\right)\ \det\left(B - XI_{n-p}\right) \\ &= \chi_{A_0}\ \det\left(B - XI_{n-p}\right) \\ &= \chi_g\ \det\left(B - XI_{n-p}\right). \end{aligned}$$
En particulier, χ_g divise χ_f. □

Lemme 1.7.5. *Soit $P = c_0 + c_1X + \cdots + X^N$ un polynôme unitaire de $\mathbb{K}[X]$ avec $N \in \mathbb{N}\backslash\{0\}$ et notons C_P la matrice*

$$C_P := \begin{pmatrix} 0 & 0 & \cdots & 0 & -c_0 \\ 1 & 0 & \cdots & 0 & -c_1 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & 1 & -c_{N-1} \end{pmatrix}$$

de $\mathrm{M}_N(\mathbb{K})$ (C_P est appelée la matrice compagnon de P). On a

$$\chi_{C_P} = (-1)^N P.$$

Démonstration. On calcule $\chi_{C_P} = \det(C_P - XI_N)$ par récurrence sur le degré N de P. Précisément, on montre que pour tout $N \in \mathbb{N}\backslash\{0\}$, pour tout polynôme unitaire P de $\mathbb{K}[X]$ de degré N , $\chi_{C_P} = (-1)^N P$.

La propriété est vraie au rang $N = 1$: soit $P = c_0 + X \in \mathbb{K}[X]$ un polynôme unitaire de degré 1, alors $C_P = (-c_0) \in \mathrm{M}_1(\mathbb{K})$ et

$$\begin{aligned} \chi_{C_P} &= \det(C_P - XI_1) \\ &= \det(-c_0 - X) \\ &= -c_0 - X \\ &= -(c_0 + X) \\ &= -P. \end{aligned}$$

Supposons à présent la propriété vraie au rang N pour $N \in \mathbb{N}\backslash\{0\}$ fixé et soit $P = c_0 + c_1X + \cdots + c_NX^N + X^{N+1} \in \mathbb{K}[X]$ un polynôme unitaire de degré $N + 1$. Alors, en utilisant un développement suivant la première ligne,

$$\begin{aligned} \chi_{C_P} &= \det(C_P - XI_{N+1}) \\ &= \det \begin{pmatrix} -X & 0 & \cdots & 0 & -c_0 \\ 1 & -X & \cdots & 0 & -c_1 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & -X & \vdots \\ 0 & \cdots & 0 & 1 & -c_N \end{pmatrix} \\ &= (-X)\det \begin{pmatrix} -X & 0 & \cdots & 0 & -c_1 \\ 1 & -X & \cdots & 0 & -c_2 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & -X & \vdots \\ 0 & \cdots & 0 & 1 & -c_N \end{pmatrix} + (-1)^{N+2}(-c_0)\det \begin{pmatrix} 1 & -X & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & -X \\ 0 & \cdots & 0 & 1 \end{pmatrix} \\ &= (-X)\chi_{C_{\tilde{P}}} + (-1)^{N+1}c_0 \end{aligned}$$

où $\widetilde{P} := c_1 + c_2X + \cdots + c_N X^{N-1} + X^N \in \mathbb{K}[X]$. En utilisant l'hypothèse de récurrence, on obtient alors

$$\begin{aligned}\chi_{C_P} &= (-X)(-1)^N \widetilde{P} + (-1)^{N+1} c_0 \\ &= (-1)^{N+1}\left(X\widetilde{P} + c_0\right) \\ &= (-1)^{N+1}\left(c_1X + c_2X^2 + \cdots + c_N X^N + X^{N+1} + c_0\right) \\ &= (-1)^{N+1} P.\end{aligned}$$

□

Démonstration du théorème 1.7.3. Afin de montrer que χ_f est un polynôme annulateur de f i.e. que $\chi_f(f) = 0_{\mathcal{L}(E)}$, nous allons montrer que pour tout vecteur $v \in E$, $\chi_f(f)(v) = 0_E$.

Soit donc v un vecteur de E, que l'on suppose de plus non nul (on a $\chi_f(f)(0_E) = 0_E$ comme $\chi_f(f)$ est une application linéaire). Soit m le plus grand entier naturel tel que la famille $\{v, f(v), \ldots, f^m(v)\}$ soit libre : un tel entier existe car la famille $\{v\}$ est libre et $m < n$ car la famille $\{v, f(v), \ldots, f^n(v)\}$ de cardinal $n+1$ est liée (l'espace vectoriel E est de dimension n). En particulier, la famille $\left\{v, f(v), \ldots, f^{m+1}(v)\right\}$ est liée : il existe $(\alpha_0, \ldots, \alpha_{m+1}) \in \mathbb{K}^{m+2} \backslash \{(0, \ldots, 0)\}$ tels que

$$\sum_{k=0}^{m+1} \alpha_k f^k(v) = 0_E.$$

De plus, $\alpha_{m+1} \neq 0$ car sinon l'égalité $\sum\limits_{k=0}^{m+1} \alpha_k f^k(v) = 0_E$ et le fait que la famille $\{v, f(v), \ldots, f^m(v)\}$ soit libre entraineraient la nullité de tous les scalaires $\alpha_0, \ldots, \alpha_m$. En posant alors, pour tout $k \in \{0, \ldots, m\}$, $a_k := -\frac{\alpha_k}{\alpha_{m+1}}$, on peut finalement écrire

$$f^{m+1}(v) = \sum_{k=0}^{m} a_k f^k(v).$$

On considère ensuite le sous-espace vectoriel $E_f := \operatorname{Vect}\{v, f(v), \ldots, f^m(v)\}$ de E engendré par la famille libre $\{v, f(v), \ldots, f^m(v)\}$: ce sous-espace vectoriel de E est stable par f car, pour tout $k \in \{0, \ldots, m\}$, $f\left(f^k(v)\right) = f^{k+1}(v) \in \operatorname{Vect}\{v, f(v), \ldots, f^m(v)\}$. On note alors $g : E_f \to E_f$; $v \mapsto f(v)$ la restriction de f à E_f.

D'après le lemme 1.7.4, χ_g divise χ_f : il existe $Q \in \mathbb{K}[X]$ tel que $\chi_f = Q\ \chi_g$ et donc $\chi_f(f)(v) = Q(f) \circ \chi_g(f)(v)$. Nous allons montrer que $\chi_g(f)(v) = 0_E$: cela entraînera alors l'égalité $\chi_f(f)(v) = 0_E$.

La matrice représentative de g dans la base $\{v, f(v), \ldots, f^m(v)\}$ de E_f est la matrice

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & a_0 \\ 1 & 0 & \cdots & 0 & a_1 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & 1 & a_m \end{pmatrix} = C_P$$

où $P := -a_0 - a_1 X - \cdots - a_m X^m + X^{m+1}$. D'où $\chi_g = \chi_{C_P} = (-1)^{m+1} P$ par le lemme précédent 1.7.5, et

$$\begin{aligned}
\chi_g(f)(v) &= \Big((-1)^{m+1} P\Big)(f)(v) \\
&= (-1)^{m+1} \left(f^{m+1} - \sum_{k=0}^{m} a_k f^k \right)(v) \\
&= (-1)^{m+1} \left(f^{m+1}(v) - \sum_{k=0}^{m} a_k f^k(v) \right) \\
&= (-1)^{m+1} \cdot 0_E \\
&= 0_E.
\end{aligned}$$

□

Ce théorème nous donne en particulier des informations supplémentaires sur μ_f :

Corollaire 1.7.6.
- *$\deg(\mu_f) \leqslant n$.*
- *Les racines de μ_f dans $\mathbb{K}$ sont exactement les racines de χ_f dans $\mathbb{K}$, i.e. les valeurs propres de f dans $\mathbb{K}$, avec multiplicités différentes a priori.*

Démonstration. Comme μ_f divise χ_f et $\deg(\chi_f) = n$ (corollaire 1.3.5), on a l'inégalité $\deg(\mu_f) \leqslant n$.

Comme μ_f divise χ_f, il existe un polynôme $P \in \mathbb{K}[X]$ tel que $\chi_f = \mu_f P$. Soit $\lambda \in \mathbb{K}$.
Si λ est une racine de μ_f, alors $\chi_f(\lambda) = \mu_f(\lambda) P(\lambda) = 0$ donc λ est une racine de χ_f (i.e. une valeur propre de f).
Réciproquement, si λ est une racine de χ_f alors λ est une valeur propre de f. Or μ_f est un polynôme annulateur de f donc, en vertu du corollaire 1.6.12, λ est une racine de P. □

Le théorème de Cayley-Hamilton nous donne ainsi un moyen de déterminer le polynôme minimal de f à partir de la donnée du polynôme caractéristique :

Exemple 1.7.7.
1. On considère la matrice
$$A = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 2 \\ 1 & 2 & 0 \end{pmatrix}$$
de $M_3(\mathbb{R})$. Son polynôme caractéristique est $\chi_A = -(X+1)(X+2)(X-3)$. Ainsi, nécessairement, $\mu_A = (X+1)(X+2)(X-3)$.
2. On considère la matrice
$$A = \begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}$$

de $M_3(\mathbb{R})$. Son polynôme caractéristique est $\chi_A = -(X-1)(X+2)^2$. Ainsi, nécessairement, $\mu_A = (X-1)(X+2)$ ou $\mu_A = (X-1)(X+2)^2$. Comme $\deg((X-1)(X+2)) < \deg\left((X-1)(X+2)^2\right)$, on commence par tester si le polynôme $(X-1)(X+2)$ annule A. On a

$$(A-I_3)(A+2I_3) = \begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

et donc $(X-1)(X+2)$ est le polynôme minimal de A.

3. On considère la matrice

$$A = \begin{pmatrix} 3 & -1 & 1 \\ 2 & 0 & 1 \\ 1 & -1 & 2 \end{pmatrix}$$

de $M_3(\mathbb{R})$. Son polynôme caractéristique est $\chi_A = -(X-1)(X-2)^2$. Ainsi, nécessairement, $\mu_A = (X-1)(X-2)$ ou $\mu_A = (X-1)(X-2)^2$. Comme $\deg((X-1)(X-2)) < \deg\left((X-1)(X-2)^2\right)$, on commence par tester si le polynôme $(X-1)(X-2)$ annule A. Or on constate que la matrice $(A-I_3)(A-2I_3)$ n'est pas la matrice nulle de $M_3(\mathbb{R})$ donc, nécessairement, $\mu_A = (X-1)(X-2)^2$.

Remarque 1.7.8. Soit A une matrice à coefficients réels. Notons $\chi_A^{\mathbb{R}}$ et $\mu_A^{\mathbb{R}}$ respectivement le polynôme caractéristique et le polynôme minimal de A en tant que matrice de $M_n(\mathbb{R})$, et $\chi_A^{\mathbb{C}}$ et $\mu_A^{\mathbb{C}}$ respectivement le polynôme caractéristique et le polynôme minimal de A en tant que matrice de $M_n(\mathbb{C})$ ($M_n(\mathbb{R}) \subset M_n(\mathbb{C})$). Alors

$$\chi_A^{\mathbb{C}} = \det(A - XI_n) = \chi_A^{\mathbb{R}} \quad \text{et} \quad \mu_A^{\mathbb{C}} = \mu_A^{\mathbb{R}}.$$

Pour établir la seconde égalité, on commence par remarquer que $\mu_A^{\mathbb{R}} \in \mathbb{R}[X] \subset \mathbb{C}[X]$ est un polynôme de $\mathbb{C}[X]$ annulant la matrice A donc $\mu_A^{\mathbb{C}}$ divise $\mu_A^{\mathbb{R}}$. Ensuite, écrivons $\mu_A^{\mathbb{C}} = P + iQ$ où $P, Q \in \mathbb{R}[X]$. On a $0_n = \mu_A^{\mathbb{C}}(A) = P(A) + iQ(A)$ donc $P(A) = 0_n$ et $Q(A) = 0_n$. Les polynômes à coefficients réels P et Q annulent la matrice à coefficients réels A, donc $\mu_A^{\mathbb{R}}$ divise les deux polynômes P et Q dans $\mathbb{R}[X]$, donc dans $\mathbb{C}[X]$, et donc $\mu_A^{\mathbb{R}}$ divise $P + iQ = \mu_A^{\mathbb{C}}$ dans $\mathbb{C}[X]$. Enfin, comme $\mu_A^{\mathbb{C}}$ et $\mu_A^{\mathbb{R}}$ sont tous deux unitaires, on obtient bien $\mu_A^{\mathbb{C}} = \mu_A^{\mathbb{R}}$.

En utilisant alors les factorisations dans $\mathbb{C}[X]$ des polynômes μ_A et χ_A, le fait qu'ils possèdent les mêmes racines dans $\mathbb{C}$ par le corollaire 1.7.6 et le fait qu'ils soient à coefficients réels, on en déduit que les polynômes μ_A et χ_A de $\mathbb{R}[X]$ possèdent les mêmes facteurs irréductibles (en général avec multiplicités différentes).

On termine cette section par un critère de diagonalisabilité permettant de décider, à partir de la donnée du polynôme minimal de f, si f est diagonalisable ou non :

Théorème 1.7.9. *L'endomorphisme f est diagonalisable si et seulement si son polynôme minimal μ_f est scindé à racines simples (dans $\mathbb{K}[X]$).*

Démonstration. Si μ_f est scindé à racines simples, alors f est diagonalisable par le théorème 1.6.15 (μ_f est, par définition, un polynôme annulateur de f).

Réciproquement, si f est diagonalisable, alors il existe, par le théorème 1.6.15, un polynôme annulateur de f qui est scindé à racines simples. Comme μ_f divise ce polynôme, μ_f est également scindé à racines simples. □

Exemple 1.7.10. Les matrices des exemples 1.7.7 1. et 2. sont diagonalisables, la matrice de l'exemple 1.7.7 3. n'est pas diagonalisable.

Remarque 1.7.11. Le théorème 1.7.9 permet de déterminer si un endomorphisme est diagonalisable ou non sans passer par le calcul des dimensions de ses espaces propres.

1.8 Réduction de Jordan des endomorphismes trigonalisables

Soit f un endomorphisme trigonalisable de E. Dans cette section, nous allons montrer, en plusieurs étapes, que f peut être réduit sous forme dite de Jordan. Plus précisément, nous allons montrer l'existence d'une base de E dans laquelle la matrice représentative de f est une matrice diagonale par blocs dont les blocs diagonaux sont des blocs de Jordan :

Définition 1.8.1. *Soit $\lambda \in \mathbb{K}$. Pour $m \in \mathbb{N}\backslash\{0\}$, on appelle λ-bloc de Jordan de taille m la matrice carrée*

$$J_m(\lambda) := \begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix}$$

de $\mathrm{M}_m(\mathbb{K})$ *; par convention,* $J_1(\lambda) := (\lambda) \in \mathrm{M}_1(\mathbb{K})$.

Pour $m_1, \ldots, m_k \in \mathbb{N}\backslash\{0\}$, on note $J_{m_1,\ldots,m_k}(\lambda)$ la matrice diagonale par blocs

$$\begin{pmatrix} J_{m_1}(\lambda) & & 0 \\ & \ddots & \\ 0 & & J_{m_k}(\lambda) \end{pmatrix}$$

de $\mathrm{M}_{m_1+\cdots+m_k}(\mathbb{K})$*. Une matrice de cette forme est dite de Jordan.*

Exemple 1.8.2. Le bloc de Jordan $J_3(-2i)$ est la matrice

$$\begin{pmatrix} -2i & 1 & 0 \\ 0 & -2i & 1 \\ 0 & 0 & -2i \end{pmatrix}$$

de $\mathrm{M}_3(\mathbb{C})$.

La matrice $J_{2,3}(1)$ est la matrice

$$\begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

de $\mathrm{M}_5(\mathbb{R})$.

Remarque 1.8.3. Pour tout $m_1, \ldots, m_k \in \mathbb{N}\backslash\{0\}$ et tout $\lambda \in \mathbb{K}$, la matrice $J_{m_1,\ldots,m_k}(\lambda)$ est triangulaire supérieure et son polynôme caractéristique est $(\lambda - X)^{m_1+\cdots+m_k}$.

Afin de pouvoir énoncer le théorème de réduction de f sous forme de Jordan, commençons par factoriser le polynôme caractéristique de f (par le théorème 1.5.3, la trigonalisabilité de f est équivalente au fait que χ_f soit scindé dans $\mathbb{K}[X]$) :

$$\chi_f = (-1)^n \prod_{i=1}^{p} (X - \lambda_i)^{m_{\lambda_i}}$$

avec $\lambda_1, \ldots, \lambda_p$ les valeurs propres deux à deux distinctes de f, et $n := \dim(E)$.

Nous avons alors le résultat de réduction suivant pour f :

Théorème 1.8.4. *Il existe une base $\mathcal{B}$ de E et, pour tout $i \in \{1, \ldots, p\}$, des entiers $m_1^i, \ldots, m_{k_i}^i \in \mathbb{N}\backslash\{0\}$ tels que*

$$\mathrm{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} J_{m_1^1,\ldots,m_{k_1}^1}(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & J_{m_1^p,\ldots,m_{k_p}^p}(\lambda_p) \end{pmatrix}.$$

De plus, à permutation près, les entiers m_j^i, $1 \leqslant i \leqslant p$, $1 \leqslant j \leqslant k_i$, sont uniques et on appelle les blocs de Jordan $J_{m_j^i}(\lambda_i)$, $1 \leqslant i \leqslant p$, $1 \leqslant j \leqslant k_i$, les blocs de Jordan de l'endomorphisme trigonalisable f. La matrice $\mathrm{Mat}_{\mathcal{B}}(f)$ est appelée la forme de Jordan de f ("la" forme de Jordan de f est unique à permutation près des blocs de Jordan).

Remarque 1.8.5.

- "La" forme de Jordan est la réduction la plus "simple" possible d'un endomorphisme trigonalisable.
- Par unicité des blocs de Jordan d'un endomorphisme trigonalisable, f est diagonalisable si et seulement si les blocs de Jordan de f sont tous de taille 1.
- Réduire sous forme de Jordan un endomorphisme trigonalisable f de E, c'est déterminer une base $\mathcal{B}$ de E dans laquelle la matrice représentative de f est de Jordan (i.e. sous la forme d'une matrice diagonale par blocs dont les blocs diagonaux sont des blocs de Jordan) et exprimer $\mathrm{Mat}_{\mathcal{B}}(f)$.
- Tout endomorphisme trigonalisable de E est réductible sous forme de Jordan.

De façon analogue, toute matrice trigonalisable A de $\mathrm{M}_n(\mathbb{K})$ est réductible sous forme de Jordan : si $\chi_A = (-1)^n \prod_{i=1}^{p} (X - \lambda_i)^{m_{\lambda_i}}$ avec $\lambda_1, \ldots, \lambda_p$ les valeurs propres deux à deux distinctes de A, alors il existe une matrice inversible $P \in \mathrm{GL}_n(\mathbb{K})$ et, pour tout $i \in \{1, \ldots, p\}$, des entiers $m_1^i, \ldots, m_{k_i}^i \in \mathbb{N}\backslash\{0\}$ tels que

$$P^{-1}AP = \begin{pmatrix} J_{m_1^1,\ldots,m_{k_1}^1}(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & J_{m_1^p,\ldots,m_{k_p}^p}(\lambda_p) \end{pmatrix}.$$

Les entiers m_j^i et les blocs de Jordan $J_{m_j^i}(\lambda_i)$, $1 \leqslant i \leqslant p$, $1 \leqslant j \leqslant k_i$ sont uniques à permutation près.

Remarquons en particulier le résultat de “classification” suivant :

Corollaire 1.8.6. *Deux matrices trigonalisables de* $\mathrm{M}_n(\mathbb{K})$ *sont semblables si et seulement si elles ont les mêmes blocs de Jordan.*

Démonstration. Soient $A, B \in \mathrm{M}_n(\mathbb{K})$ et supposons qu'il existe $P, Q \in \mathrm{GL}_n(\mathbb{K})$ telles que

$$P^{-1}AP = \begin{pmatrix} J_1 & & 0 \\ & \ddots & \\ 0 & & J_r \end{pmatrix} = Q^{-1}BQ$$

avec $J_1, \dots, J_r$ des blocs de Jordan, alors, en particulier, $P^{-1}AP = Q^{-1}BQ$ et donc

$$\left(PQ^{-1}\right)^{-1} A \left(PQ^{-1}\right) = B.$$

□

Remarque 1.8.7. <u>Réduire sous forme de Jordan</u> une matrice trigonalisable $A \in \mathrm{M}_n(\mathbb{K})$, c'est déterminer une matrice inversible $P \in \mathrm{GL}_n(\mathbb{K})$ telle que la matrice $P^{-1}AP$ soit de Jordan et exprimer $P^{-1}AP$.

Nous allons montrer une partie du théorème 1.8.4, à savoir l'existence d'une réduction de Jordan pour l'endomorphisme trigonalisable f, ceci en détaillant une méthode systématique de réduction sous forme de Jordan. On résumera ensuite les étapes-clés de cette procédure dans le langage des matrices.

Dans ce document, nous ne montrerons pas l'unicité, à permutation près, des blocs de Jordan de f.

La première étape pour réduire sous forme de Jordan l'endomorphisme trigonalisable f est de considérer ses sous-espaces caractéristiques.

1.8.1 Réduction suivant les sous-espaces caractéristiques

Rappelons encore une fois que l'on a supposé dans cette section que l'endomorphisme f était trigonalisable.

Définition 1.8.8. *Soit* $\lambda \in \mathrm{Sp}(f)$. *On appelle <u>sous-espace caractéristique de f associé à la valeur propre λ</u> le sous-espace vectoriel*

$$N_\lambda := \mathrm{Ker}\ \left(f - \lambda \mathrm{Id}_E\right)^{m_\lambda}$$

de E.

Remarque 1.8.9. Avec les notations ci-dessus,

- $E_\lambda \subset N_\lambda$, car si $v \in E$ vérifie $(f - \lambda \mathrm{Id}_E)(v) = 0_E$, alors
$$(f - \lambda \mathrm{Id}_E)^{m_\lambda}(v) = (f - \lambda \mathrm{Id}_E)^{m_\lambda - 1}\big((f - \lambda \mathrm{Id}_E)(v)\big) = (f - \lambda \mathrm{Id}_E)^{m_\lambda - 1}(0_E) = 0_E,$$
- $f(N_\lambda) \subset N_\lambda$ (autrement dit N_λ est stable par f), car si $v \in N_\lambda = \mathrm{Ker}\ (f - \lambda \mathrm{Id}_E)^{m_\lambda}$, alors
$$\begin{aligned}(f - \lambda \mathrm{Id}_E)^{m_\lambda}(f(v)) &= (X - \lambda)^{m_\lambda}(f) \circ X(f)(v)\\ &= X(f) \circ (X - \lambda)^{m_\lambda}(f)(v)\\ &= f\big((f - \lambda \mathrm{Id}_E)^{m_\lambda}(v)\big)\\ &= f(0_E)\\ &= 0_E.\end{aligned}$$

Exemple 1.8.10. Reprenons la matrice
$$A = \begin{pmatrix} 3 & -1 & 1 \\ 2 & 0 & 1 \\ 1 & -1 & 2 \end{pmatrix} \in \mathrm{M}_2(\mathbb{R})$$
de l'exemple 1.5.5. Son polynôme caractéristique est $\chi_A = -(X-1)(X-2)^2$. Le sous-espace caractéristique de A associé à la valeur propre 1 est $N_1 = \mathrm{Ker}\ (A - I_3) = E_1 = \mathrm{Vect}\left\{\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}\right\}$ et le sous-espace caractéristique de A associé à la valeur propre 2 est $N_2 = \mathrm{Ker}\ (A - 2I_3)^2$. Or
$$(A - 2I_3)^2 = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 1 & 0 \end{pmatrix}$$
donc $N_2 = \mathrm{Ker}\ (A - 2I_3)^2 = \mathrm{Vect}\left\{\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right\}$.

Une propriété remarquable des sous-espaces caractéristiques d'un endomorphisme trigonalisable est qu'ils sont en somme directe et que leur somme est égale à l'espace E tout entier :

Proposition 1.8.11. *Les sous-espaces caractéristiques de f sont en somme directe et, reprenant les notations du début de la section,*
$$E = N_{\lambda_1} \oplus \cdots \oplus N_{\lambda_p}.$$

Cette proposition est une conséquence du lemme des noyaux :

Théorème 1.8.12 (Lemme des noyaux). *Soit h un endomorphisme de E et soient $P_1, \ldots, P_r$ des polynômes de $\mathbb{K}[X]$ premiers entre eux deux à deux. Alors les sous-espaces vectoriels* $\mathrm{Ker}\ P_i(h)$, *$i \in \{1, \ldots, r\}$, de E sont en somme directe et*
$$\bigoplus_{i=1}^{r} \mathrm{Ker}\ P_i(h) = \mathrm{Ker}\ \left(\prod_{i=1}^{r} P_i\right)(h).$$

Démonstration. On montre tout d'abord le résultat pour deux polynômes P et Q de $\mathbb{K}[X]$ premiers entre eux, puis on montrera le résultat général par récurrence sur r.

On a $(PQ)(h) = P(h) \circ Q(h) = Q(h) \circ P(h)$ donc Ker $P(h) \subset$ Ker $(PQ)(h)$ et Ker $Q(h) \subset$ Ker $(PQ)(h)$ donc

$$\text{Ker } P(h) + \text{Ker } Q(h) \subset \text{Ker } (PQ)(h).$$

Réciproquement, montrons que Ker $(PQ)(h) \subset$ Ker $P(h)$ + Ker $Q(h)$. Comme P et Q sont premiers entre eux, il existe une relation de Bézout $UP + VQ = 1$ avec $U, V \in \mathbb{K}[X]$, et alors $U(h) \circ P(h) + V(h) \circ Q(h) = \mathrm{Id}_E$. Soit maintenant $v \in$ Ker $(PQ)(h)$, d'après l'égalité précédente,

$$v = \mathrm{Id}_E(v) = U(h) \circ P(h)(v) + V(h) \circ Q(h)(v).$$

Or

$$Q(h)\Big(U(h) \circ P(h)(v)\Big) = U(h)\Big((PQ)(h)(v)\Big) = U(h)\,(0_E) = 0_E$$

et de la même façon $P(h)\Big(V(h) \circ Q(h)(v)\Big) = 0_E$, donc $v \in$ Ker $P(h)$+Ker $Q(h)$. En conclusion,

$$\text{Ker } (PQ)(h) = \text{Ker } P(h) + \text{Ker } Q(h).$$

Enfin, montrons que cette dernière somme est directe : soit $v \in$ Ker $P(h) \cap$ Ker $Q(h)$, alors

$$\begin{aligned} v &= U(h) \circ P(h)(v) + V(h) \circ Q(h)(v) \\ &= U(h)\Big(P(h)(v)\Big) + V(h)\Big(Q(h)(v)\Big) \\ &= U(h)\,(0_E) + V(h)\,(0_E) \\ &= 0_E. \end{aligned}$$

A présent, montrons par récurrence sur $r \in \mathbb{N}\backslash\{0\}$ que pour tout $r \in \mathbb{N}\backslash\{0\}$, pour tous polynômes $P_1, \dots, P_r \in \mathbb{K}[X]$ premiers entre eux deux à deux, les sous-espaces vectoriels Ker $P_i(h)$, $i \in \{1, \dots, r\}$, de E sont en somme directe et que

$$\bigoplus_{i=1}^{r} \text{Ker } P_i(h) = \text{Ker } \left(\prod_{i=1}^{r} P_i\right)(h).$$

La propriété est vraie au rang $r = 1$.

Supposons maintenant la propriété vraie au rang r pour $r \in \mathbb{N}\backslash\{0\}$ fixé et soient $P_1, \dots, P_{r+1}$ des polynômes de $\mathbb{K}[X]$ premiers entre eux deux à deux. Comme le polynôme P_{r+1} est premier avec chacun des polynômes $P_1, \dots, P_r$, il est premier avec le produit $\prod_{i=1}^{r} P_i$. D'après l'hypothèse de récurrence, $\bigoplus_{i=1}^{r} \text{Ker } P_i(h) = \text{Ker } \left(\prod_{i=1}^{r} P_i\right)(h)$ et, d'après ce qui a été démontré ci-dessus, les espaces Ker $P_{r+1}(h)$ et Ker $\left(\prod_{i=1}^{r} P_i\right)(h) = \bigoplus_{i=1}^{r} \text{Ker } P_i(h)$ sont en somme directe : les espaces Ker $P_i(h)$, $i \in \{1, \dots, r+1\}$ sont donc en somme directe. Ensuite,

$$\begin{aligned} \text{Ker } P_{r+1}(h) \oplus \left(\bigoplus_{i=1}^{r} \text{Ker } P_i(h)\right) &= \text{Ker } P_{r+1}(h) \oplus \text{Ker } \left(\prod_{i=1}^{r} P_i\right)(f) \\ &= \text{Ker } \left(P_{r+1} \left(\prod_{i=1}^{r} P_i\right)(f)\right) \end{aligned}$$

donc

$$\bigoplus_{i=1}^{r+1} \text{Ker } P_i(h) = \text{Ker } \left(\prod_{i=1}^{r+1} P_i\right)(f).$$

□

Remarque 1.8.13. Dans la preuve ci-dessus, nous avons utilisé le fait suivant : pour $r \in \mathbb{N}\backslash\{0\}$, si $E_1, \dots, E_{r+1}$ sont des sous-espaces vectoriels de E tels que les espaces $E_1, \dots, E_r$ sont en somme directe et E_{r+1} est en somme directe avec la somme $\bigoplus_{i=1}^{r} E_i$, alors les espaces $E_1, \dots, E_r, E_{r+1}$ sont en somme directe. En effet, soient $v_1, \dots, v_{r+1} \in E$ tels que, pour tout $i \in \{1, \dots, r+1\}$, $v_i \in E_i$, et $v_1 + \dots + v_r + v_{r+1} = 0_E$. Alors $(v_1 + \dots + v_r) + v_{r+1} = 0_E$ donc, comme $\bigoplus_{i=1}^{r} E_i$ et E_{r+1} sont en somme directe, $v_1 + \dots + v_r = 0_E$ et $v_{r+1} = 0_E$. Enfin, comme les espaces $E_1, \dots, E_r$ sont en somme directe, l'égalité $v_1 + \dots + v_r = 0_E$ implique la nullité des vecteurs $v_1, \dots, v_r$.

Démonstration de la proposition 1.8.11. Les valeurs propres $\lambda_1, \dots, \lambda_p$ de f étant deux à deux distinctes, les polynômes $(X - \lambda_i)^{m_{\lambda_i}}$, $i \in \{1, \dots, p\}$, sont premiers entre eux deux à deux. D'après le lemme des noyaux, les sous-espaces caractéristiques $N_{\lambda_i} = \text{Ker } (f - \lambda_i \text{Id}_E)^{m_{\lambda_i}} = \text{Ker } (X - \lambda_i \text{Id}_E)^{m_{\lambda_i}}(f)$, $i \in \{1, \dots, p\}$, de f sont donc en somme directe et

$$N_{\lambda_1} \oplus \cdots \oplus N_{\lambda_p} = \text{Ker } \left(\prod_{i=1}^{p} (X - \lambda_i)^{m_{\lambda_i}}\right)(f).$$

Mais $\prod_{i=1}^{p} (X - \lambda_i)^{m_{\lambda_i}} = (-1)^n \chi_f$ et $\chi_f(f) = 0_{\mathcal{L}(E)}$ donc Ker $\left(\prod_{i=1}^{p} (X - \lambda_i)^{m_{\lambda_i}}\right)(f) = E$ et

$$N_{\lambda_1} \oplus \cdots \oplus N_{\lambda_p} = E.$$

□

Une conséquence importante de la proposition 1.8.11 et du fait que chaque sous-espace caractéristique N_{λ_i}, $i \in \{1, \dots, p\}$, de f soit stable par f (cf. remarque 1.8.9) est la suivante.

Pour $i \in \{1, \ldots, p\}$, soit $\mathcal{B}_i$ une base de N_{λ_i} et notons $\mathcal{B}_0 := \{\mathcal{B}_1, \ldots, \mathcal{B}_p\}$. On a alors, comme $E = N_{\lambda_1} \oplus \cdots \oplus N_{\lambda_p}$ (proposition 1.8.11) et comme chaque sous-espace caractéristique de f est stable par f (remarque 1.8.9),

$$\mathrm{Mat}_{\mathcal{B}_0}(f) = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_p \end{pmatrix}$$

où, pour tout $i \in \{1, \ldots, p\}$, $A_i := \mathrm{Mat}_{\mathcal{B}_i}\left(f_{|N_{\lambda_i}}\right)$ (cette réduction est une “diagonalisation par blocs” de f).

Remarquons alors que si, pour tout $i \in \{1, \ldots, p\}$, on montre l'existence d'une base $\mathcal{B}'_i$ de N_{λ_i} telle que $\mathrm{Mat}_{\mathcal{B}'_i}\left(f_{|N_{\lambda_i}}\right) = J_{m^i_1, \ldots, m^i_{k_i}}(\lambda_i)$ avec $m^i_1, \ldots, m^i_{k_i} \in \mathbb{N}\backslash\{0\}$, alors, en notant $\mathcal{B} := \{\mathcal{B}'_1, \ldots, \mathcal{B}'_p\}$, on obtiendra la réduction

$$\mathrm{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} J_{m^1_1, \ldots, m^1_{k_1}}(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & J_{m^p_1, \ldots, m^p_{k_p}}(\lambda_p) \end{pmatrix},$$

qui est la réduction sous forme de Jordan recherchée.

Pour montrer l'existence de telles familles $\mathcal{B}'_i$, $i \in \{1, \ldots, p\}$, et les déterminer, on commence par écrire, pour $i \in \{1, \ldots, p\}$,

$$f_{N_{\lambda_i}} = \lambda_i \mathrm{Id}_{N_{\lambda_i}} + (f - \lambda_i \mathrm{Id}_E)_{|N_{\lambda_i}},$$

puis on utilise le fait que l'endomorphisme $(f - \lambda_i \mathrm{Id}_E)_{|N_{\lambda_i}} = f_{|N_{\lambda_i}} - \lambda_i \mathrm{Id}_{|N_{\lambda_i}}$ de N_{λ_i} est nilpotent :

1.8.2 Réduction de Jordan des endomorphismes nilpotents

Commençons par introduire la notion d'endomorphisme nilpotent et de matrice nilpotente :

Définition 1.8.14. *Soit u un endomorphisme de E. On dit que u est nilpotent s'il existe $l \in \mathbb{N}\backslash\{0\}$ tel que $u^l = 0_{\mathcal{L}(E)}$. Dans ce cas, on appelle indice de nilpotence de u le plus petit entier naturel non nul $\nu \in \mathbb{N}\backslash\{0\}$ tel que $u^\nu = 0_{\mathcal{L}(E)}$.*

De façon analogue, une matrice carrée $U \in \mathrm{M}_n(\mathbb{K})$ est dite nilpotente s'il existe $l \in \mathbb{N}\backslash\{0\}$ tel que $U^l = 0_n$ et, dans ce cas, on appelle indice de nilpotence de A le plus petit entier naturel non nul $\nu \in \mathbb{N}\backslash\{0\}$ tel que $A^\nu = 0_n$.

Remarque 1.8.15.
- Un endomorphisme u de $\mathcal{L}(E)$ est nilpotent si et seulement s'il existe $l \in \mathbb{N}\backslash\{0\}$ tel que le polynôme X^l annule u.
- L'indice de nilpotence d'un endomorphisme nilpotent ν est l'unique entier naturel non nul ν tel que $u^{\nu-1} \neq 0_{\mathcal{L}(E)}$ et $u^\nu = 0_{\mathcal{L}(E)}$ (car si $u^{l_0} = 0_{\mathcal{L}(E)}$ pour $l_0 \in \mathbb{N}\backslash\{0\}$, alors $u^l = 0_{\mathcal{L}(E)}$ pour tout $l \in \mathbb{N}\backslash\{0\}$ supérieur ou égal à l_0).

- Si u est un endomorphisme nilpotent d'indice de nilpotence ν, alors $\mu_u = X^\nu$. En effet, le polynôme X^ν annule l'endomorphisme u donc μ_u divise X^ν, donc $\mu_u = X^{l_0}$ avec $l_0 \in \{1, \ldots, \nu\}$, mais, pour tout $l \in \{1, \ldots, \nu - 1\}$, $X^l(u) = u^l \neq 0_{\mathcal{L}(E)}$ par définition de l'indice de nilpotence, donc, nécessairement, $l_0 = \nu$.
- On a les propriétés analogues sur les matrices nilpotentes.
- Si $\mathcal{B}$ est une base de E, alors un endomorphisme u de E est nilpotent si et seulement si sa matrice représentative $\mathrm{Mat}_{\mathcal{B}}(u)$ dans la base $\mathcal{B}$ est nilpotente.

Exemple 1.8.16.
- La matrice
$$U := \begin{pmatrix} -1 & 0 & 1 & 0 \\ -3 & -2 & 5 & 0 \\ -2 & -1 & 3 & 0 \\ 2 & 1 & -2 & 0 \end{pmatrix}$$
de $\mathrm{M}_4(\mathbb{R})$ est nilpotente d'indice de nilpotence 4, car
$$U^2 = \begin{pmatrix} -1 & -1 & 2 & 0 \\ -1 & -1 & 2 & 0 \\ -1 & -1 & 2 & 0 \\ -1 & 0 & 1 & 0 \end{pmatrix}, \ U^3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & -1 & 2 & 0 \end{pmatrix}, \ U^4 = 0_4.$$
- La matrice
$$U := \begin{pmatrix} 5 & -3 & 2 \\ 15 & -9 & 6 \\ 10 & -6 & 4 \end{pmatrix}$$
de $\mathrm{M}_3(\mathbb{R})$ est nilpotente d'indice de nilpotence 2 car $U^2 = 0_3$.
- La matrice
$$U := \begin{pmatrix} 2 & 0 & 4 & -2 & -3 \\ -2 & 0 & -3 & 2 & 4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 2 & -1 & -2 \end{pmatrix}$$
de $\mathrm{M}_5(\mathbb{R})$ est nilpotente d'indice de nilpotence 3 car
$$U^2 = \begin{pmatrix} 1 & 0 & 2 & -1 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 2 & -1 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \ U^3 = 0_5.$$

Traitons deux autres exemples :

Lemme 1.8.17. *Les matrices triangulaires de* $\mathrm{M}_n(\mathbb{K})$ *dont les coefficients diagonaux sont tous nuls sont nilpotentes, d'indice de nilpotence inférieur ou égal à* n.

Démonstration. Soit U une matrice triangulaire de $\mathrm{M}_n(\mathbb{K})$ dont tous les coefficients diagonaux sont nuls. On suppose tout d'abord que U est triangulaire supérieure. Nous allons montrer par récurrence sur $k \in \mathbb{N}\backslash\{0\}$ que pour tout $k \in \mathbb{N}\backslash\{0\}$, pour tous $i, j \in \{1, \ldots, n\}$ tels que $j < i + k$, le coefficient situé à la ligne i et la colonne j de la matrice U^k est nulle. En particulier, $U^n = 0_n$ et l'indice de nilpotence de U est donc inférieur ou égal à n.

La propriété est vraie au rang $k = 1$ par hypothèse sur U. Supposons maintenant la propriété vraie au rang k pour $k \in \mathbb{N}\backslash\{0\}$ fixé, et soient $i, j \in \{1, \ldots, n\}$ tels que $j < i + k + 1$. Si on note $U^k = (a_{r\,s})_{1 \leqslant r,s \leqslant n}$ et $U = (b_{r\,s})_{1 \leqslant r,s \leqslant n}$, le coefficient situé à la ligne i et la colonne j de la matrice $U^{k+1} = U^k\, U$ est

$$\begin{aligned} \sum_{t=1}^{n} a_{i,t} b_{t,j} &= \sum_{t=i+k}^{n} a_{i,t} b_{t,j} \text{ (car } a_{i,t} = 0 \text{ si } t < i + k) \\ &= \sum_{i+k \leqslant t \leqslant j-1} a_{i,t} b_{t,j} \text{ (car } b_{t,j} = 0 \text{ si } j < t + 1 \Leftrightarrow t > j - 1) \\ &= 0 \text{ (la somme ci-dessus est vide car } j < i + k + 1 \Leftrightarrow i + k > j - 1). \end{aligned}$$

Si U est une matrice triangulaire inférieure, alors sa transposée ${}^tU$ est une matrice triangulaire supérieure dont tous les coefficients diagonaux sont nuls et on a alors

$$U^n = {}^t\left({}^tU\right)^n = {}^t0_n = 0_n.$$

□

Exemple 1.8.18. Pour tous $m_1, \ldots, m_k \in \mathbb{N}\backslash\{0\}$, la matrice de Jordan $J_{m_1,\ldots,m_k}(0)$ est nilpotente.

Lemme 1.8.19. *Reprenant les notations de la partie précédente, si $\lambda \in \mathrm{Sp}(f)$, l'endomorphisme $(f - \lambda \mathrm{Id}_E)_{|N_\lambda} = f_{|N_\lambda} - \lambda \mathrm{Id}_{|N_\lambda}$ de N_λ est nilpotent, d'indice de nilpotence inférieur ou égal à m_λ.*

Démonstration. Par définition de $N_\lambda = \mathrm{Ker}\ (f - \lambda \mathrm{Id}_E)^{m_\lambda}$, tout vecteur v de N_λ vérifie

$$\left(f_{|N_\lambda} - \lambda \mathrm{Id}_{|N_\lambda}\right)^{m_\lambda}(v) = (f - \lambda \mathrm{Id}_E)^{m_\lambda}(v) = 0_E = 0_{N_\lambda},$$

i.e. $\left(f_{|N_\lambda} - \lambda \mathrm{Id}_{|N_\lambda}\right)^{m_\lambda} = 0_{\mathcal{L}(N_\lambda)}$. □

Nous allons montrer que tout endomorphisme nilpotent – et toute matrice nilpotente – est réductible sous forme de Jordan. Soit u un endomorphisme nilpotent de E d'ordre de nilpotence $\nu \in \mathbb{N}\backslash\{0\}$. Nous allons construire, à l'aide d'un procédé algorithmique, une base de E dans laquelle la matrice représentative de u est de Jordan :

Théorème 1.8.20. *Il existe une base $\mathcal{B}$ de E et des entiers $m_1, \ldots, m_k \in \mathbb{N}\backslash\{0\}$ tels que*

$$\mathrm{Mat}_{\mathcal{B}}(u) = J_{m_1,\ldots,m_k}(0) = \begin{pmatrix} J_{m_1}(0) & & 0 \\ & \ddots & \\ 0 & & J_{m_k}(0) \end{pmatrix}.$$

En particulier, u est trigonalisable et $\chi_u = (-1)^n X^n$.

Démonstration. Pour $r \in \{0, \dots, \nu\}$, notons $M_r := \operatorname{Ker} u^r$. Remarquons que l'on a alors une suite d'inclusions

$$\{0_E\} = M_0 \subset M_1 \subset \cdots \subset M_{\nu-1} \subset M_\nu = E$$

et que, si $r \in \{1, \dots, \nu\}$, $u(M_r) \subset M_{r-1}$ (car si $v \in M_r = \operatorname{Ker} u^r$, $u^{r-1}\big(u(v)\big) = u^r(v) = 0_E$).

Par récurrence descendante sur $r \in \{1, \dots, \nu\}$, nous allons construire des sous-espaces vectoriels $S_1, \dots, S_\nu$ de E tels que

- pour tout $r \in \{1, \dots, \nu\}$, $S_r \oplus M_{r-1} = M_r$ (i.e. S_r est un supplémentaire de M_{r-1} dans M_r),
- tout $r \in \{1, \dots, \nu-1\}$, $u(S_{r+1}) \subset S_r$ et la restriction $u_{|S_{r+1}} : S_{r+1} \to S_r \subset E$ est injective.

Des bases bien choisies des espaces $S_1, \dots, S_\nu$ nous fourniront ensuite une base de E dans laquelle la matrice représentative de u est de la forme voulue.

Le point de départ de cette construction est de considérer un supplémentaire S_ν de $M_{\nu-1}$ dans $M_\nu = E$. Supposons ensuite que, pour $r \in \{2, \dots, \nu\}$ fixé, on ait déjà construit les espaces $S_r, \dots, S_\nu$ vérifiant les propriétés voulues. En particulier, $S_r \subset M_r$ donc $u(S_r) \subset u(M_r) \subset M_{r-1}$. On a alors :

- L'image $u(S_r)$ est en somme directe avec M_{r-2} (car si $u(v) \in M_{r-2} = \operatorname{Ker} u^{r-2}$ avec $v \in S_r$, on a $u^{r-1}(v) = 0_E$ donc $v \in S_r \cap M_{r-1} = \{0_E\}$, donc $u(v) = 0_E$) et on peut donc construire un supplémentaire S_{r-1} de M_{r-2} dans M_{r-1} contenant $u(S_r)$: précisément, on considère une base $\mathcal{F}$ de $u(S_r)$ et une base $\mathcal{G}$ de M_{r-2}, on complète la famille libre $\mathcal{F} \sqcup \mathcal{G}$ de M_{r-1} en une base $\mathcal{F} \sqcup \mathcal{G} \sqcup \widetilde{\mathcal{F}}$ de M_{r-1}, puis on pose $S_{r-1} := \operatorname{Vect}\Big(\mathcal{F} \sqcup \widetilde{\mathcal{F}}\Big)$),
- la restriction $u_{|S_r} : S_r \to u(S_r) \subset S_{r-1} \subset E$ est injective (car si $v \in S_r$ vérifie $u(v) = 0_E$, alors $v \in \operatorname{Ker} u \subset \operatorname{Ker} u^{r-1} = M_{r-1}$, or $S_r \cap M_{r-1} = \{0_E\}$ donc $v = 0_E$).

Une fois cette construction achevée, on a alors

$$\begin{aligned}
E &= M_\nu \\
&= M_{\nu-1} \oplus S_\nu \\
&= (M_{\nu-2} \oplus S_{\nu-1}) \oplus S_\nu \\
&\cdots \\
&= (M_0 \oplus S_1) \oplus S_2 \oplus \cdots \oplus S_\nu \\
&= S_1 \oplus \cdots \oplus S_\nu
\end{aligned}$$

(on utilise ici la propriété énoncée dans la remarque 1.8.13).

Ensuite, soit $\mathcal{B}_\nu$ une base de S_ν et, pour tout $r \in \{1, \dots, \nu-1\}$, soit $\mathcal{B}_r$ une base de S_r telle que la famille $\mathcal{B}_r$ contient la famille $u(\mathcal{B}_{r+1})$ (i.e. la famille libre constituée des images par u des vecteurs de la famille $\mathcal{B}_{r+1}$).

La réunion $\mathcal{B}$ des familles libres $\mathcal{B}_1, \dots, \mathcal{B}_\nu$ de E est une base de E (car $E = S_1 \oplus \cdots \oplus S_\nu$) et

- si $v \in \mathcal{B}_1 \subset M_1 = \operatorname{Ker} u$, $u(v) = 0$,

- si $v \in \mathcal{B}_{r+1}$ avec $r \in \{1, \ldots, \nu - 1\}$, il existe un unique vecteur $w \in \mathcal{B}_r$ tel que $u(v) = w$.

Ainsi, en réordonnant convenablement les vecteurs de $\mathcal{B}$, on peut écrire

$$\mathcal{B} = \{v_1^1, \ldots, v_{m_1}^1, \ldots, v_1^k, \ldots, v_{m_k}^k\}$$

de manière que

- pour tout $i \in \{1, \ldots, k\}$, $u\left(v_1^i\right) = 0_E$,
- pour tout $i \in \{1, \ldots, k\}$ et tout $s \in \{2, \ldots, m_i\}$, $u\left(v_s^i\right) = v_{s-1}^i$,

et la matrice représentative de u dans $\mathcal{B}$ est alors

$$\mathrm{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} J_{m_1}(0) & & 0 \\ & \ddots & \\ 0 & & J_{m_k}(0) \end{pmatrix} = J_{m_1, \ldots, m_k}(0).$$

□

La preuve précédente fournit <u>une</u> méthode constructive de réduction sous forme de Jordan des endomorphismes nilpotents, basée sur le calcul de noyaux et la complétion de familles libres en bases. Rappelons que, pour compléter une famille libre en une base, on peut utiliser le résultat suivant : si $\mathcal{B}$ est une base de E et si $\mathcal{C}$ est une famille libre de E, il existe (au moins) une façon de compléter la famille libre $\mathcal{C}$ de E en une base $\mathcal{C} \sqcup \mathcal{B}'$ de E en utilisant des vecteurs de la base $\mathcal{B}$ (i.e. de manière que $\mathcal{B}' \subset \mathcal{B}$), et alors, si on note $F := \mathrm{Vect}\, \mathcal{C}$ et $F' := \mathrm{Vect}\, \mathcal{B}'$, F' est un supplémentaire de F dans E.

Illustrons cette procédure avec quelques exemples matriciels :

Exemple 1.8.21. On réduit sous forme de Jordan les matrices de l'exemple 1.8.16.

- On considère la matrice

$$U := \begin{pmatrix} -1 & 0 & 1 & 0 \\ -3 & -2 & 5 & 0 \\ -2 & -1 & 3 & 0 \\ 2 & 1 & -2 & 0 \end{pmatrix}$$

de $\mathrm{M}_4(\mathbb{R})$, nilpotente d'indice 4.

<u>Construction de "$\mathcal{B}_4$" :</u> On a

$$M_4 := \mathrm{Ker}\ U^4 = \mathrm{M}_{4,1}(\mathbb{R}) = \mathrm{Vect} \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

et

$$M_3 := \mathrm{Ker}\ U^3 = \mathrm{Ker} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & -1 & 2 & 0 \end{pmatrix} = \mathrm{Vect} \left\{ \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

On note $\mathcal{C}_4 := \left\{ \begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\1\end{pmatrix} \right\}$ (c'est une base de M_4) et $\mathcal{C}_3 := \left\{ \begin{pmatrix}1\\-1\\0\\0\end{pmatrix}, \begin{pmatrix}1\\1\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\1\end{pmatrix} \right\}$ (c'est une base de M_3), et on choisit d'utiliser le vecteur $\begin{pmatrix}1\\0\\0\\0\end{pmatrix}$ de $\mathcal{C}_4$ pour compléter la famille libre $\mathcal{C}_3$ de M_4 en une base de M_4 : on pose alors $\mathcal{B}_4 := \left\{ \begin{pmatrix}1\\0\\0\\0\end{pmatrix} \right\}$ et $S_4 := \operatorname{Vect} \mathcal{B}_4$.

Construction de "$\mathcal{B}_3$" : On a

$$M_2 := \operatorname{Ker} U^2 = \operatorname{Ker} \begin{pmatrix} -1 & -1 & 2 & 0 \\ -1 & -1 & 2 & 0 \\ -1 & -1 & 2 & 0 \\ -1 & 0 & 1 & 0 \end{pmatrix} = \operatorname{Vect} \left\{ \begin{pmatrix}1\\1\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\1\end{pmatrix} \right\}.$$

et on note $\mathcal{C}_2 := \left\{ \begin{pmatrix}1\\1\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\1\end{pmatrix} \right\}$ (c'est une base de M_2). On a $U(\mathcal{B}_4) = \left\{ \begin{pmatrix}-1\\-3\\-2\\2\end{pmatrix} \right\}$ et on pose $\mathcal{B}_3 := \left\{ \begin{pmatrix}-1\\-3\\-2\\2\end{pmatrix} \right\}$ et $S_3 := \operatorname{Vect} \mathcal{B}_3 = U(S_4)$ (l'image $U(S_4)$ est déjà un supplémentaire de M_2 dans M_3).

Construction de "$\mathcal{B}_2$" : On a

$$M_1 := \operatorname{Ker} U = \operatorname{Ker} \begin{pmatrix} -1 & 0 & 1 & 0 \\ -3 & -2 & 5 & 0 \\ -2 & -1 & 3 & 0 \\ 2 & 1 & -2 & 0 \end{pmatrix} = \operatorname{Vect} \left\{ \begin{pmatrix}0\\0\\0\\1\end{pmatrix} \right\},$$

et on note $\mathcal{C}_1 := \left\{ \begin{pmatrix}0\\0\\0\\1\end{pmatrix} \right\}$ (c'est une base de M_1). On a $U(\mathcal{B}_3) = \left\{ \begin{pmatrix}-1\\-1\\-1\\-1\end{pmatrix} \right\}$ et on pose $\mathcal{B}_2 := \left\{ \begin{pmatrix}-1\\-1\\-1\\-1\end{pmatrix} \right\}$ et $S_2 := \operatorname{Vect} \mathcal{B}_2 = U(S_3)$ (l'image $U(S_3)$ est déjà un supplémentaire de M_1 dans M_2).

Construction de "$\mathcal{B}_1$" : On a

$$M_0 := \text{Ker } I_n = \left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\}.$$

On a $U(\mathcal{B}_2) = \left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \\ -1 \end{pmatrix} \right\}$ et on pose $\mathcal{B}_1 := \left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \\ -1 \end{pmatrix} \right\}$ et $S_1 := \operatorname{Vect} \mathcal{B}_1 = U(S_2) = M_1$ (l'image $U(S_2)$ est déjà égale à M_1).

Construction de "$\mathcal{B}$" : Si on écrit alors

$$\mathcal{B} := \left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ -3 \\ -2 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\},$$

la famille $\mathcal{B}$ est une base de $\mathrm{M}_{4,1}(\mathbb{R})$ et, en posant

$$P := \begin{pmatrix} 0 & -1 & -1 & 1 \\ 0 & -1 & -3 & 0 \\ 0 & -1 & -2 & 0 \\ -1 & -1 & 2 & 0 \end{pmatrix},$$

on a

$$P^{-1}UP = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

- On considère maintenant la matrice

$$U := \begin{pmatrix} 5 & -3 & 2 \\ 15 & -9 & 6 \\ 10 & -6 & 4 \end{pmatrix}$$

de $\mathrm{M}_3(\mathbb{R})$, nilpotente d'indice 2.

Construction de "$\mathcal{B}_2$" : On a

$$M_2 := M_{3,1}(\mathbb{R}) = \operatorname{Vect} \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

et

$$M_1 := \operatorname{Ker} U = \operatorname{Vect}\left\{\begin{pmatrix}1\\1\\-1\end{pmatrix}, \begin{pmatrix}0\\2\\3\end{pmatrix}\right\}.$$

On note $\mathcal{C}_2 := \left\{\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}\right\}$ (c'est une base de M_2) et $\mathcal{C}_1 := \left\{\begin{pmatrix}1\\1\\-1\end{pmatrix}, \begin{pmatrix}0\\2\\3\end{pmatrix}\right\}$ (c'est une base de M_1), et on <u>choisit</u> d'utiliser le vecteur $\begin{pmatrix}0\\0\\1\end{pmatrix}$ de $\mathcal{C}_2$ pour compléter la famille libre $\mathcal{C}_1$ de M_2 en une base de M_2 : on pose alors $\mathcal{B}_2 := \left\{\begin{pmatrix}0\\0\\1\end{pmatrix}\right\}$ et $S_2 := \operatorname{Vect} \mathcal{B}_2$.

<u>Construction de "$\mathcal{B}_1$" :</u> On a

$$M_0 := \left\{\begin{pmatrix}0\\0\\0\end{pmatrix}\right\}.$$

On a $U(\mathcal{B}_2) = \left\{\begin{pmatrix}2\\6\\4\end{pmatrix}\right\}$, et on <u>choisit</u> d'utiliser le vecteur $\begin{pmatrix}1\\1\\-1\end{pmatrix}$ de $\mathcal{C}_1$ pour compléter la famille libre $\left\{\begin{pmatrix}2\\6\\4\end{pmatrix}\right\}$ de M_1 en une base de M_1 : on pose alors $\mathcal{B}_1 := \left\{\begin{pmatrix}2\\6\\4\end{pmatrix}, \begin{pmatrix}1\\1\\-1\end{pmatrix}\right\}$ et $S_1 := \operatorname{Vect} \mathcal{B}_1 = M_1$.

<u>Construction de "$\mathcal{B}$" :</u> Si on écrit alors

$$\mathcal{B} := \left\{\begin{pmatrix}2\\6\\4\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}, \begin{pmatrix}1\\1\\-1\end{pmatrix}\right\},$$

la famille $\mathcal{B}$ est une base de $\mathrm{M}_{3,1}(\mathbb{R})$ et, en posant

$$P := \begin{pmatrix}2 & 0 & 1\\6 & 0 & 1\\4 & 1 & -1\end{pmatrix},$$

on a

$$P^{-1}UP = \begin{pmatrix}0 & 1 & 0\\0 & 0 & 0\\0 & 0 & 0\end{pmatrix}.$$

- Terminons ces exemples avec la matrice

$$U := \begin{pmatrix} 2 & 0 & 4 & -2 & -3 \\ -2 & 0 & -3 & 2 & 4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 2 & -1 & -2 \end{pmatrix}$$

de $\mathrm{M}_5(\mathbb{R})$, nilpotente d'indice 3.

<u>Construction de "$\mathcal{B}_3$" :</u> On a

$$M_3 := M_{5,1}(\mathbb{R}) = \mathrm{Vect}\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

et

$$M_2 := \mathrm{Ker}\, U^2 = \mathrm{Ker} \begin{pmatrix} 1 & 0 & 2 & -1 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 2 & -1 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = \mathrm{Vect}\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right\}.$$

On note $\mathcal{C}_3$, resp. $\mathcal{C}_2$, la base de M_3, resp. M_2, considérée ci-dessus, et on choisit d'utiliser le vecteur $\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ de $\mathcal{C}_3$ pour compléter la famille libre $\mathcal{C}_2$ de M_3 en une base de M_3 : on pose alors $\mathcal{B}_3 := \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\}$.

<u>Construction de "$\mathcal{B}_2$" :</u> On a

$$M_1 := \mathrm{Ker}\, U = \mathrm{Ker} \begin{pmatrix} 2 & 0 & 4 & -2 & -3 \\ -2 & 0 & -3 & 2 & 4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 2 & -1 & -2 \end{pmatrix} = \mathrm{Vect}\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

On note $\mathcal{C}_1$ la base ci-dessus considérée de M_1, et on a $U(\mathcal{B}_3) = \begin{pmatrix} 2 \\ -2 \\ 0 \\ 0 \\ 1 \end{pmatrix}$: on utilise alors le vecteur $\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}$ de $\mathcal{C}_2$ pour compléter la famille libre $\mathcal{C}_1 \sqcup \left\{ \begin{pmatrix} 2 \\ -2 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}$ de M_2 en une base de M_2 : on pose ensuite $\mathcal{B}_2 := \left\{ \begin{pmatrix} 2 \\ -2 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} \right\}$.

<u>Construction de "$\mathcal{B}_1$" :</u> On a $U(\mathcal{B}_2) = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\}$, qui est une base de M_1, et on pose $\mathcal{B}_1 := U(\mathcal{B}_2)$

<u>Construction de "$\mathcal{B}$" :</u> Si on écrit alors

$$\mathcal{B} := \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ -2 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} \right\},$$

la famille $\mathcal{B}$ est une base de $\mathrm{M}_{5,1}(\mathbb{R})$ et, en posant

$$P := \begin{pmatrix} 1 & 2 & 1 & 1 & 0 \\ 0 & -2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{pmatrix},$$

on a

$$P^{-1}UP = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Remarque 1.8.22. Reprenant les notations de la preuve du théorème 1.8.20, si l'indice de nilpotence ν de l'endomorphisme nilpotent u est égal à la dimension n de E, pour tout $r \in \{1, \ldots, \nu\}$, la base $\mathcal{B}_r$ de S_r est constituée d'un seul vecteur (car la réunion des n familles libres disjointes $\mathcal{B}_1, \ldots, \mathcal{B}_n$ est une base de l'espace vectoriel E de dimension n) et donc, en particulier, pour tout $r \in \{1, \ldots, n-1\}$, $\mathcal{B}_r = u\left(\mathcal{B}_{r+1}\right)$.

Ainsi, si $\nu = n$, pour construire la base "$\mathcal{B}$", il suffit de choisir tout d'abord un vecteur v de E qui ne soit pas dans $M_{n-1} = \operatorname{Ker} u^{n-1}$, puis de considérer ses images successives par les composées itérées de u : la matrice représentative de u dans la base $\mathcal{B} = \left\{u^{n-1}(v), u^{n-2}(v), \ldots, u(v), v\right\}$ est alors $J_n(0)$.

1.8.3 Preuve de la réductibilité sous forme de Jordan des endomorphismes trigonalisables

Reprenons les notations du début de la section : f désigne un endomorphisme trigonalisable de E, $\chi_f = (-1)^n \prod_{i=1}^{p} (X - \lambda_i)^{m_{\lambda_i}}$ avec $\lambda_1, \ldots, \lambda_p$ les valeurs propres deux à deux distinctes de f et on a montré, dans la proposition 1.8.11, que E était égal à la somme directe des espaces caractéristiques $N_{\lambda_1}, \ldots, N_{\lambda_p}$ de f.

En conséquence, si, pour $i \in \{1, \ldots, p\}$, $\mathcal{B}_i$ désigne une base de N_{λ_i} et $\mathcal{B}_0 := \{\mathcal{B}_1, \ldots, \mathcal{B}_p\}$, on a

$$\operatorname{Mat}_{\mathcal{B}_0}(f) = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_p \end{pmatrix}$$

où, pour tout $i \in \{1, \ldots, p\}$, $A_i := \operatorname{Mat}_{\mathcal{B}_i}\left(f_{|N_{\lambda_i}}\right)$.

Pour terminer la preuve de l'existence d'une réduction de Jordan pour f, nous allons montrer, pour tout $i \in \{1, \ldots, p\}$, l'existence d'une base $\mathcal{B}'_i$ de N_{λ_i} dans laquelle la matrice de $f_{|N_{\lambda_i}} \in \mathcal{L}\left(N_{\lambda_i}\right)$ est de Jordan, ceci en appliquant le théorème 1.8.20 à l'endomorphisme nilpotent $f_{|N_\lambda} - \lambda \mathrm{Id}_{|N_\lambda}$ de N_λ.

À un certain point, nous évoquerons le fait que la dimension du sous-espace caractéristique associé à une valeur propre λ de f est sa multiplicité m_λ. Nous le montrons dans le lemme suivant. Avant de l'énoncer, remarquons que

$$\begin{aligned}
\chi_f &= \det\left(\operatorname{Mat}_{\mathcal{B}_0}(f - X\mathrm{Id}_E)\right) \\
&= \det\left(\operatorname{Mat}_{\mathcal{B}_0}(f)\right) - XI_n \\
&= \det \begin{pmatrix} A_1 - XI_{n_1} & & 0 \\ & \ddots & \\ 0 & & A_p - XI_{n_p} \end{pmatrix} \text{ (où, pour } i \in \{1, \ldots, p\}\text{, on note } n_i := \dim\left(N_{\lambda_i}\right)\text{)}, \\
&= \det\left(A_1 - XI_{n_1}\right) \cdots \det\left(A_p - XI_{n_p}\right) \\
&= \chi_{f_{|N_{\lambda_1}}} \cdots \chi_{f_{|N_{\lambda_p}}}
\end{aligned}$$

Lemme 1.8.23. *Soit $\lambda \in \operatorname{Sp}(f)$ et notons $N := N_\lambda$.*

1. *La seule valeur propre de la restriction $f_{|N} \in \mathcal{L}(N)$ est λ.*

2. $\chi_{f_{|N}} = (\lambda - X)^{m_\lambda}$.

3. $\dim(N) = m_\lambda$.

Démonstration. 1. Soit $\lambda_0 \in \mathbb{K}$ une valeur propre de $f_{|N}$ et soit v un vecteur propre associé (en particulier, $v \neq 0_N$). Alors $f_{|N}(v) - \lambda v = \lambda_0 v - \lambda v = (\lambda_0 - \lambda)(v)$: le vecteur v est donc un vecteur propre pour la valeur propre $\lambda_0 - \lambda$ de $f_{|N} - \lambda \mathrm{Id}_N$. Mais ce dernier endomorphisme de N est nilpotent (cf. lemme 1.8.19) et la seule valeur propre d'un endomorphisme nilpotent est 0 (par le théorème 1.8.20), donc $\lambda_0 - \lambda = 0$ i.e. $\lambda_0 = \lambda$.

2. Comme le sous-espace vectoriel N de E est stable par $f \in \mathcal{L}(E)$ (cf. remarque 1.8.9), par le lemme 1.7.4, le polynôme $\chi_{f_{|N}}$ divise $\chi_f = (-1)^n \prod_{i=1}^{p} (X - \lambda_i)^{m_{\lambda_i}}$. Comme la seule valeur propre de f est λ, nécessairement $\chi_{f_{|N}} = (-1)^l (X - \lambda)^l$ avec $l \in \{1, \ldots, m_\lambda\}$.
Ainsi, pour tout $i \in \{1, \ldots, p\}$, il existe $l_i \in \{1, \ldots, m_{\lambda_i}\}$ tel que $\chi_{f_{|N_{\lambda_i}}} = (-1)^{l_i}(X - \lambda_i)^{l_i}$. Mais $\chi_f = \chi_{f_{N_{\lambda_1}}} \cdots \chi_{f_{N_{\lambda_p}}}$ donc, nécessairement, pour tout $i \in \{1, \ldots, p\}$, $\chi_{f_{|N_{\lambda_i}}} = (-1)^{m_{\lambda_i}}(X - \lambda_i)^{m_{\lambda_i}}$. En particulier, $\chi_{f_{|N}} = (-1)^{m_\lambda}(X - \lambda)^{m_\lambda}$.

3. Le degré du polynôme caractéristique d'un endomorphisme d'un espace vectoriel est égal à la dimension de ce dernier (cf. corollaire 1.3.5) et $\deg\left(\chi_{f_{|N}}\right) = m_\lambda$ donc $\dim(N) = m_\lambda$.
□

Démonstration du théorème 1.8.4 (preuve de l'existence d'une réduction de Jordan pour f). Soit $\lambda \in \mathrm{Sp}(f)$ et, pour simplifier les écritures, notons $N := N_\lambda$ et $m := m_\lambda = \dim N$. Nous allons montrer qu'il existe une base $\mathcal{B}'$ de N et des entiers $m_1, \ldots, m_k \in \mathbb{N}\backslash\{0\}$ tels que $\mathrm{Mat}_{\mathcal{B}'}\left(f_{|N}\right) = J_{m_1,\ldots,m_k}(\lambda)$.

Pour ce faire, on écrit
$$f_{|N} = \lambda \mathrm{Id}_N + f_{|N} - \lambda \mathrm{Id}_N$$
et on utilise la nilpotence de l'endomorphisme $f_{|N} - \lambda \mathrm{Id}_N$ de N (cf. lemme 1.8.19) : d'après le théorème 1.8.20, il existe une base $\mathcal{B}'$ de N et des entiers $m_1, \ldots, m_k \in \mathbb{N}\backslash\{0\}$ tels que $\mathrm{Mat}_{\mathcal{B}'}\left(f_{|N} - \lambda \mathrm{Id}_N\right) = J_{m_1,\ldots,m_k}(0)$, et alors
$$\begin{aligned}
\mathrm{Mat}_{\mathcal{B}'}\left(f_{|N}\right) &= \mathrm{Mat}_{\mathcal{B}'}\left(\lambda \mathrm{Id}_N + f_{|N} - \lambda \mathrm{Id}_N\right) \\
&= \mathrm{Mat}_{\mathcal{B}'}\left(\lambda \mathrm{Id}_N\right) + \mathrm{Mat}_{\mathcal{B}'}\left(f_{|N} - \lambda \mathrm{Id}_N\right) \\
&= \lambda I_m + J_{m_1,\ldots,m_k}(0) \\
&= J_{m_1,\ldots,m_k}(\lambda).
\end{aligned}$$

Ainsi, nous avons bien montré que pour tout $i \in \{1, \ldots, p\}$, il existe une base $\mathcal{B}'_i$ de N_{λ_i} et des entiers $m_1^i, \ldots, m_{k_i}^i \in \mathbb{N}\backslash\{0\}$ tels que $\mathrm{Mat}_{\mathcal{B}'_i}\left(f_{|N_{\lambda_i}}\right) = J_{m_1^i,\ldots,m_{k_i}^i}(\lambda_i)$: si l'on note alors $\mathcal{B} := \left\{\mathcal{B}'_1, \ldots, \mathcal{B}'_p\right\}$, on a
$$\mathrm{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} J_{m_1^1,\ldots,m_{k_1}^1}(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & J_{m_1^p,\ldots,m_{k_p}^p}(\lambda_p) \end{pmatrix}.$$

□

1.8.4 Description matricielle de la méthode de réduction sous forme de Jordan

Pour finir, résumons la méthode de réduction sous forme de Jordan proposée dans cette section, dans sa version matricielle. Soit donc $A \in \mathrm{M}_n(\mathbb{K})$ une matrice trigonalisable. On suppose que l'on a déjà écrit le polynôme caractéristique χ_A de A comme un produit

$$\chi_A = (-1)^n \prod_{i=1}^{p} (X - \lambda_i)^{m_{\lambda_i}}$$

dans $\mathbb{K}[X]$, où $\lambda_1, \dots, \lambda_p$ sont les valeurs propres deux à deux distinctes de A.

Étape 1 : Pour tout $i \in \{1, \dots, p\}$, calculer la matrice $(A - \lambda_i I_n)^{m_{\lambda_i}}$ et déterminer une base $\mathcal{B}_i$ de $N_{\lambda_i} = \mathrm{Ker}\ (A - \lambda_i I_n)^{m_{\lambda_i}} \subset \mathrm{M}_{n,1}(\mathbb{K})$. Considérer la base $\mathcal{B}_0 := \{\mathcal{B}_1, \dots, \mathcal{B}_p\}$ de $\mathrm{M}_{n,1}(\mathbb{K})$ et la matrice P_0 dont les colonnes sont, dans l'ordre, les vecteurs colonnes de la base $\mathcal{B}_0$. Calculer la matrice $P_0^{-1} A P_0$: elle est de la forme

$$\begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_p \end{pmatrix}$$

où, pour tout $i \in \{1, \dots, p\}$, $A_i \in \mathrm{M}_{m_{\lambda_i}}(\mathbb{K})$ (et $\chi_{A_i} = (-1)^{m_{\lambda_i}} (X - \lambda_i)^{m_{\lambda_i}}$).

Étape 2 : Pour chaque $i \in \{1, \dots, p\}$, calculer la matrice $U_i := A_i - \lambda_i I_{m_{\lambda_i}}$ de $M_{m_{\lambda_i}}(\mathbb{K})$ puis appliquer à la matrice nilpotente U_i la méthode de réduction des matrices nilpotentes à la forme de Jordan décrite dans la preuve du théorème 1.8.20 : on obtient des entiers $m_1^i, \dots, m_{k_i}^i \in \mathbb{N} \backslash \{0\}$ et une matrice inversible Q_i de taille m_{λ_i} tels que $Q_i^{-1} U_i Q_i = J_{m_1^i, \dots, m_{k_i}^i}(0)$.

Étape 3 : On note

$$\widetilde{P} := \begin{pmatrix} Q_1 & & 0 \\ & \ddots & \\ 0 & & Q_p \end{pmatrix} \in \mathrm{GL}_n(\mathbb{K})$$

et alors

$$\begin{aligned}
\widetilde{P}^{-1} \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_p \end{pmatrix} \widetilde{P} &= \begin{pmatrix} \lambda_1 I_{m_{\lambda_1}} & & 0 \\ & \ddots & \\ 0 & & \lambda_p I_{m_{\lambda_p}} \end{pmatrix} + \begin{pmatrix} J_{m_1^1, \dots, m_{k_1}^1}(0) & & 0 \\ & \ddots & \\ 0 & & J_{m_1^p, \dots, m_{k_p}^p}(0) \end{pmatrix} \\
&= \begin{pmatrix} J_{m_1^1, \dots, m_{k_1}^1}(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & J_{m_1^p, \dots, m_{k_p}^p}(\lambda_p) \end{pmatrix}.
\end{aligned}$$

En posant $P := P_0\widetilde{P}$, on obtient donc

$$P^{-1}AP = \begin{pmatrix} J_{m_1^1,\ldots,m_{k_1}^1}(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & J_{m_1^p,\ldots,m_{k_p}^p}(\lambda_p) \end{pmatrix}.$$

Appliquons cette méthode à quelques exemples :

Exemple 1.8.24. On considère la matrice

$$A := \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 4 & 1 & -2 \\ 2 & 1 & 2 & -1 \\ 1 & 2 & 1 & 0 \end{pmatrix}$$

de $\mathrm{M}_4(\mathbb{R})$.

<u>Étape 0 :</u> On calcule le polynôme caractéristique χ_A de A :

$$\begin{aligned}
\chi_A = \det(A - XI_4) \quad &= \begin{vmatrix} 1-X & 0 & 0 & 0 \\ -1 & 4-X & 1 & -2 \\ 2 & 1 & 2-X & -1 \\ 1 & 2 & 1 & -X \end{vmatrix} \\
&= (1-X)\begin{vmatrix} 4-X & 1 & -2 \\ 1 & 2-X & -1 \\ 2 & 1 & -X \end{vmatrix} \\
&\underset{C_1 \leftarrow C_1 + C_3}{=} (1-X)\begin{vmatrix} 2-X & 1 & -2 \\ 0 & 2-X & -1 \\ 2-X & 1 & -X \end{vmatrix} \\
&= (1-X)(2-X)\begin{vmatrix} 1 & 1 & -2 \\ 0 & 2-X & -1 \\ 1 & 1 & -X \end{vmatrix} \\
&\underset{L3 \leftarrow L_3 - L_1}{=} (1-X)(2-X)\begin{vmatrix} 1 & 1 & -2 \\ 0 & 2-X & -1 \\ 0 & 0 & 2-X \end{vmatrix} \\
&= (1-X)(2-X)^3
\end{aligned}$$

<u>Étape 1 :</u> La multiplicité de la valeur propre 1 dans χ_A étant 1, $N_1 = E_1$, et

$$E_1 = \mathrm{Ker}\ (A - I_4) = \mathrm{Ker}\ \begin{pmatrix} 0 & 0 & 0 & 0 \\ -1 & 3 & 1 & -2 \\ 2 & 1 & 1 & -1 \\ 1 & 2 & 1 & -1 \end{pmatrix} = \mathrm{Vect}\left\{\begin{pmatrix} 1 \\ 1 \\ -4 \\ -1 \end{pmatrix}\right\}.$$

La multiplicité de la valeur propre 2 dans χ_A est 3. Pour déterminer $N_2 = \text{Ker } (A-2I_4)^3$, on commence par calculer la matrice $(A-2I_4)^3$. On a

$$(A-2I_4)^3 = \begin{pmatrix} -1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 4 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

et donc

$$N_2 = \text{Vect}\left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

On note P_0 la matrice inversible

$$\begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -4 \\ 0 & 0 & 1 & -1 \end{pmatrix}$$

et on obtient

$$P_0^{-1}AP_0 = \begin{pmatrix} 4 & 1 & -2 & 0 \\ 1 & 2 & -1 & 0 \\ 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$$

Étape 2 : Le bloc $A_2 = (1) \in \mathrm{M}_1(\mathbb{R})$ est déjà un bloc de Jordan $J_1(1)$.

On note U_1 la matrice

$$A_1 - 2I_3 = \begin{pmatrix} 4 & 1 & -2 \\ 1 & 2 & -1 \\ 2 & 1 & 0 \end{pmatrix} - 2I_3 = \begin{pmatrix} 2 & 1 & -2 \\ 1 & 0 & -1 \\ 2 & 1 & -2 \end{pmatrix} \in M_3(\mathbb{R})$$

et on applique la méthode de réduction à la forme de Jordan des matrices nilpotentes à U_1 :

Étape a : Déterminons l'indice de nilpotence de U_1 : on a

$$U_1^2 = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{pmatrix}$$

et U_1^3 est la matrice nulle de $\mathrm{M}_3(\mathbb{R})$. Ainsi, l'indice de nilpotence de U_1 est 3.

Étape b : Comme l'indice de nilpotence de U_1 est égal à la dimension de l'espace $M_{3,1}(\mathbb{R})$ (i.e la multiplicité de la valeur propre 2 dans χ_A), on choisit ensuite un vecteur colonne qui n'est pas dans le noyau de U_1^2 : on prend par exemple le vecteur colonne $Y := \begin{pmatrix}1\\0\\0\end{pmatrix}$, on calcule $U_1 Y = \begin{pmatrix}2\\1\\2\end{pmatrix}$ puis $U_1^2 Y = \begin{pmatrix}1\\0\\1\end{pmatrix}$, et la famille libre $\{U_1^2 Y, U_1 Y, Y\}$ est une base de $\mathrm{M}_{3,1}(\mathbb{R})$. Si on pose

$$Q_1 := \begin{pmatrix}1 & 2 & 1\\0 & 1 & 0\\1 & 2 & 0\end{pmatrix},$$

on a ainsi

$$Q_1^{-1} U_1 Q = \begin{pmatrix}0 & 1 & 0\\0 & 0 & 1\\0 & 0 & 0\end{pmatrix}$$

(cf. remarque 1.8.22).

Etape 3 : On note

$$\widetilde{P} := \begin{pmatrix}Q_1 & 0\\0 & 1\end{pmatrix} = \begin{pmatrix}1 & 2 & 1 & 0\\0 & 1 & 0 & 0\\1 & 2 & 0 & 0\\0 & 0 & 0 & 1\end{pmatrix}$$

et

$$P := P_0 \widetilde{P} = \begin{pmatrix}0 & 0 & 0 & 1\\1 & 2 & 1 & 1\\0 & 1 & 0 & -4\\1 & 2 & 0 & -1\end{pmatrix},$$

et on a

$$P^{-1} A P = \begin{pmatrix}2 & 1 & 0 & 0\\0 & 2 & 1 & 0\\0 & 0 & 2 & 0\\0 & 0 & 0 & 1\end{pmatrix} = \begin{pmatrix}J_3(2) & 0\\0 & J_1(1)\end{pmatrix}.$$

Exemple 1.8.25. On considère la matrice

$$A := \begin{pmatrix}5 & 0 & 4 & -2 & -3\\-2 & 3 & -3 & 2 & 4\\0 & 0 & 3 & 0 & 0\\0 & 0 & 0 & 3 & 1\\1 & 0 & 2 & -1 & 1\end{pmatrix}$$

de $\mathrm{M}_5(\mathbb{R})$.

Étape 0 : Calculons le polynôme caractéristique χ_A de A :

$$\begin{aligned}
\chi_A = \det(A - XI_5) &= \begin{vmatrix} 5-X & 0 & 4 & -2 & -3 \\ -2 & 3-X & -3 & 2 & 4 \\ 0 & 0 & 3-X & 0 & 0 \\ 0 & 0 & 0 & 3-X & 1 \\ 1 & 0 & 2 & -1 & 1-X \end{vmatrix} \\
&= (3-X)\begin{vmatrix} 5-X & 0 & -2 & -3 \\ -2 & 3-X & 2 & 4 \\ 0 & 0 & 3-X & 1 \\ 1 & 0 & -1 & 1-X \end{vmatrix} \\
&= (3-X)^2\begin{vmatrix} 5-X & -2 & -3 \\ 0 & 3-X & 1 \\ 1 & -1 & 1-X \end{vmatrix} \\
&\underset{C_1 \leftarrow C_1 + C_2}{=} (3-X)^2\begin{vmatrix} 3-X & -2 & -3 \\ 3-X & 3-X & 1 \\ 0 & -1 & 1-X \end{vmatrix} \\
&= (3-X)^3\begin{vmatrix} 1 & -2 & -3 \\ 1 & 3-X & 1 \\ 0 & -1 & 1-X \end{vmatrix} \\
&\underset{L_2 \leftarrow L_2 - L_1}{=} (3-X)^3\begin{vmatrix} 1 & -2 & -3 \\ 0 & 5-X & 4 \\ 0 & -1 & 1-X \end{vmatrix} \\
&= (3-X)^3\begin{vmatrix} 5-X & 4 \\ -1 & 1-X \end{vmatrix} \\
&= (3-X)^3\left[(5-X)(1-X)+4\right] \\
&= (3-X)^3(X^2-6X+9) \\
&= (3-X)^5
\end{aligned}$$

3 est donc l'unique valeur propre de A et la matrice

$$A - 3I_5 = \begin{pmatrix} 2 & 0 & 4 & -2 & -3 \\ -2 & 0 & -3 & 2 & 4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 2 & -1 & -2 \end{pmatrix}$$

est nilpotente.

On peut alors directement passer à l'étape 2 de la méthode et remarquer que la matrice

$U := A - 3I_5$ a déjà été réduite sous forme de Jordan dans l'exemple 1.8.21 : en posant

$$P := \begin{pmatrix} 1 & 2 & 1 & 1 & 0 \\ 0 & -2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{pmatrix},$$

on a

$$P^{-1}UP = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

et donc

$$P^{-1}AP = \begin{pmatrix} 3 & 1 & 0 & 0 & 0 \\ 0 & 3 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix} = J_{3,2}(3).$$

1.9 Décomposition de Dunford des endomorphismes trigonalisables

Au cours de notre réduction sous forme de Jordan des endomorphismes et matrices trigonalisables, nous avons (quasiment) démontré le théorème de "décomposition" suivant. Une fois n'est pas coutume, nous commencerons par énoncer et démontrer (du moins en partie) le résultat dans le langage des matrices :

Théorème 1.9.1. *Soit $A \in \mathrm{M}_n(\mathbb{K})$ une matrice trigonalisable. il existe une unique matrice diagonalisable $D \in \mathrm{M}_n(\mathbb{K})$ et une unique matrice nilpotente $U \in \mathrm{M}_n(\mathbb{K})$ telles que $DU = UD$ et $A = D + U$.*

L'écriture $A = D+U$, avec D et U comme ci-dessus, est appelée la décomposition de Dunford de A.

Nous allons montrer l'existence d'une décomposition de Dunford pour $A \in \mathrm{M}_n(\mathbb{K})$ trigonalisable en réduisant A suivant ses sous-espaces caractéristiques (section 1.8.1 ou "étape 1" de la méthode de réduction sous forme de Jordan). Dans ce document, nous ne montrerons en revanche pas l'unicité de la décomposition de Dunford de A.

Démonstration du théorème 1.9.1 (preuve de l'existence d'une décomposition de Dunford). On considère l'écriture du polynôme caractéristique χ_A de A comme un produit

$$\chi_A = (-1)^n \prod_{i=1}^{p} (X - \lambda_i)^{m_{\lambda_i}}$$

dans $\mathbb{K}[X]$, où $\lambda_1, \ldots, \lambda_p$ sont les valeurs propres deux à deux distinctes de A. Nous avons montré qu'il existait une matrice inversible $P_0 \in \mathrm{GL}_n(\mathbb{K})$ et, pour tout $i \in \{1, \ldots, p\}$, une matrice $A_i \in \mathrm{M}_{m_{\lambda_i}}(\mathbb{K})$ telles que

$$P_0^{-1} A P_0 = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_p \end{pmatrix}$$

et, pour tout $i \in \{1, \ldots, p\}$, la matrice $U_i := A_i - \lambda_i I_{m_{\lambda_i}}$ de $\mathrm{M}_{m_{\lambda_i}}(\mathbb{K})$ est nilpotente.

On écrit alors

$$P_0^{-1} A P_0 = \begin{pmatrix} \lambda_1 I_{m_{\lambda_1}} & & 0 \\ & \ddots & \\ 0 & & \lambda_p I_{m_{\lambda_p}} \end{pmatrix} + \begin{pmatrix} U_1 & & 0 \\ & \ddots & \\ 0 & & U_p \end{pmatrix}$$

et on pose $D_0 := \begin{pmatrix} \lambda_1 I_{m_{\lambda_1}} & & 0 \\ & \ddots & \\ 0 & & \lambda_p I_{m_{\lambda_p}} \end{pmatrix}$ et $U_0 := \begin{pmatrix} U_1 & & 0 \\ & \ddots & \\ 0 & & U_p \end{pmatrix}$.

On a :

- D_0 est une matrice diagonale,
- U_0 est une matrice nilpotente, car si $m := \max\limits_{1 \leqslant i \leqslant p} m_{\lambda_i}$,

$$U_0^m = \begin{pmatrix} U_1^m & & 0 \\ & \ddots & \\ 0 & & U_p^m \end{pmatrix} = \begin{pmatrix} 0_{m_{\lambda_1}} & & 0 \\ & \ddots & \\ 0 & & 0_{m_{\lambda_p}} \end{pmatrix} = 0_n,$$

- D_0 et U_0 commutent (i.e. $D_0 U_0 = U_0 D_0$), car

$$D_0 U_0 = \begin{pmatrix} \left(\lambda_1 I_{m_{\lambda_1}}\right) U_1 & & 0 \\ & \ddots & \\ 0 & & \left(\lambda_p I_{m_{\lambda_p}}\right) U_p \end{pmatrix} = \begin{pmatrix} U_1 \left(\lambda_1 I_{m_{\lambda_1}}\right) & & 0 \\ & \ddots & \\ 0 & & U_p \left(\lambda_p I_{m_{\lambda_p}}\right) \end{pmatrix} = U_0 D_0,$$

- $P_0^{-1} A P_0 = D_0 + U_0$.

On écrit enfin

$$A = P_0 D_0 P_0^{-1} + P_0 U_0 P_0^{-1}$$

et on pose $D := P_0 D_0 P_0^{-1}$ et $U := P_0 U_0 P_0^{-1}$.

Alors :

- D est une matrice diagonalisable (car $P_0^{-1} D P_0 = D_0$ est une matrice diagonale),

- U est une matrice nilpotente, car, avec les notations ci-dessus,

$$U^m = \left(P_0 U_0 P_0^{-1}\right)^m = P_0 U_0^m P_0^{-1} = 0_n,$$

- D et U commutent, car

$$\begin{aligned} DU &= \left(P_0 D_0 P_0^{-1}\right)\left(P_0 U_0 P_0^{-1}\right) \\ &= P_0 \left(D_0 U_0\right) P_0^{-1} \\ &= P_0 \left(U_0 D_0\right) P_0^{-1} \\ &= \left(P_0 U_0 P_0^{-1}\right)\left(P_0 D_0 P_0^{-1}\right) \\ &= UD, \end{aligned}$$

- $A = D + U$.

□

Exemple 1.9.2. On calcule la décomposition de Dunford de la matrice

$$A := \begin{pmatrix} 2 & 1 & -1 \\ 3 & 3 & -4 \\ 3 & 1 & -2 \end{pmatrix}$$

de $\mathrm{M}_3(\mathbb{R})$.

On a $\chi_A = (2-X)^2(-1-X)$, $N_2 = \mathrm{Ker}\ (A-2I_3)^2 = \mathrm{Ker}\ \begin{pmatrix} 0 & 0 & 0 \\ -9 & 0 & 9 \\ -9 & 0 & 9 \end{pmatrix} = \mathrm{Vect}\left\{\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right\}$ et $N_{-1} = E_{-1} = \mathrm{Ker}\ (A + I_3) = \mathrm{Ker}\ \begin{pmatrix} 3 & 1 & -1 \\ 3 & 4 & -4 \\ 3 & 1 & -1 \end{pmatrix} = \mathrm{Vect}\left\{\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}\right\}$.

On pose ensuite $P := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \in \mathrm{GL}_n(\mathbb{R})$ et, sachant que $P^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & -1 \\ -1 & 0 & 1 \end{pmatrix}$,

$$D := P \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix} P^{-1} = \begin{pmatrix} 2 & 0 & 0 \\ 3 & 2 & -3 \\ 3 & 0 & -1 \end{pmatrix}.$$

On pose enfin

$$U := A - D = \begin{pmatrix} 0 & 1 & -1 \\ 0 & 1 & -1 \\ 0 & 1 & -1 \end{pmatrix},$$

et alors $A = D + U$ est la décomposition de Dunford de A.

Remarque 1.9.3. On peut également obtenir la décomposition de Dunford d'une matrice trigonalisable $A \in \mathrm{M}_n(\mathbb{K})$ à partir d'une réduction sous forme de Jordan de A, en suivant le même principe que dans la preuve du théorème 1.9.1. Cependant, la réduction sous forme de Jordan est plus coûteuse en calculs que la réduction suivant les sous-espaces caractéristiques (d'autant plus si, comme dans la méthode présentée dans la section 1.8, on utilise la réduction suivant les sous-espaces caractéristiques pour réduire sous forme de Jordan).

Énonçons maintenant le théorème de décomposition de Dunford pour les endomorphismes trigonalisables de E.

Théorème 1.9.4. *Si f est un endomorphisme trigonalisable de E, il existe un unique endomorphisme diagonalisable d de $\mathcal{L}(E)$ et un unique endomorphisme nilpotent u de $\mathcal{L}(E)$ tels que $d \circ u = u \circ d$ et $f = d + u$.*

L'écriture $f = d + u$, avec d et u comme ci-dessus, est appelée la décomposition de Dunford de f.

Démonstration. Soit $\mathcal{B}$ une base de E et notons $A := \mathrm{Mat}_{\mathcal{B}}(f)$. D'après le théorème 1.9.1, il existe une matrice diagonalisable $D \in \mathrm{M}_n(\mathbb{K})$ et une matrice nilpotente $U \in \mathrm{M}_n(\mathbb{K})$ tels que $DU = UD$ et $A = D + U$.

Notons d, resp. u, l'endomorphisme de E dont la matrice représentative dans la base $\mathcal{B}$ est D, resp. U. Alors :

- d est diagonalisable car la matrice D est diagonalisable (si la matrice $P^{-1}DP$ est diagonale avec $P \in \mathrm{GL}_n(\mathbb{K})$, alors P peut être considérée comme la matrice de passage de la base $\mathcal{B}$ à une base $\mathcal{B}'$ de E et $\mathrm{Mat}_{\mathcal{B}'}(d) = P^{-1}DP$),
- u est nilpotent car la matrice U est nilpotente (cf. remarque 1.8.15),
- les endomorphismes d et u commutent car leurs matrices représentatives, respectivement D et U, dans la base $\mathcal{B}$ commutent,
- $f = d + u$ car $A = D + U$.

Montrons maintenant l'unicité de cette décomposition : soient $\widetilde{d}$ un endomorphisme diagonalisable de E et $\widetilde{u}$ un endomorphisme nilpotent de E tels que $\widetilde{d} \circ \widetilde{u} = \widetilde{u} \circ \widetilde{d}$ et $f = \widetilde{d} + \widetilde{u}$, et notons $\widetilde{D} := \mathrm{Mat}_{\mathcal{B}}\left(\widetilde{d}\right) \in \mathrm{M}_n(\mathbb{K})$ et $\widetilde{U} := \mathrm{Mat}_{\mathcal{B}}(\widetilde{u}) \in \mathrm{M}_n(\mathbb{K})$. On a alors $A = \widetilde{D} + \widetilde{U}$ avec $\widetilde{D}$ diagonalisable et $\widetilde{U}$ nilpotente telles que $\widetilde{D}\widetilde{U} = \widetilde{U}\widetilde{D}$: par unicité de la décomposition de Dunford de A, on a donc $\widetilde{D} = D$ et $\widetilde{U} = U$, i.e. $\mathrm{Mat}_{\mathcal{B}}\left(\widetilde{d}\right) = \mathrm{Mat}_{\mathcal{B}}(d)$ et $\mathrm{Mat}_{\mathcal{B}}(\widetilde{u}) = \mathrm{Mat}_{\mathcal{B}}(u)$, i.e. $\widetilde{d} = d$ et $\widetilde{u} = u$. $\square$

La décomposition de Dunford permet entre autres choses de calculer les puissances successives d'une matrice trigonalisable. Précisément, soit A une matrice trigonalisable de $\mathrm{M}_n(\mathbb{K})$ de décomposition de Dunford $A = D + U$ avec D diagonalisable et U nilpotente. Alors, comme D et U commutent, on peut calculer $A^k = (D + U)^k$, pour tout $k \in \mathbb{N}$, à l'aide du binôme de Newton. De plus, la nilpotence de U simplifie l'expression du développement (en tout cas lorsque k est plus grand que l'indice de nilpotence de U).

Exemple 1.9.5. On cherche à calculer les puissances successives de la matrice

$$A := \begin{pmatrix} 2 & 1 & -1 \\ 3 & 3 & -4 \\ 3 & 1 & -2 \end{pmatrix} \in \mathrm{M}_3(\mathbb{R})$$

de l'exemple 1.9.2. On avait déterminé la décomposition de Dunford $A = D + U$ de A avec $D := \begin{pmatrix} 2 & 0 & 0 \\ 3 & 2 & -3 \\ 3 & 0 & -1 \end{pmatrix}$ diagonalisable et $U := \begin{pmatrix} 0 & 1 & -1 \\ 0 & 1 & -1 \\ 0 & 1 & -1 \end{pmatrix}$ nilpotente d'indice de nilpotence égal à 2. D'où, si $k \in \mathbb{N}\backslash\{0\}$,

$$\begin{aligned} A^k &= (D+U)^k \\ &= \sum_{i=0}^{k} \binom{k}{i} D^{k-i} U^i \text{ (les matrices } D \text{ et } U \text{ commutent)} \\ &= \sum_{i=0}^{1} \binom{k}{i} D^{k-i} U^i \ (U^i = 0_n \text{ pour tout } i \geqslant 2) \\ &= D^k + kD^{k-1}U \end{aligned}$$

Or $D = P \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix} P^{-1}$ avec $P := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$, donc

$$D^k = P \begin{pmatrix} 2^k & 0 & 0 \\ 0 & 2^k & 0 \\ 0 & 0 & (-1)^k \end{pmatrix} P^{-1} = \begin{pmatrix} 2^k & 0 & 0 \\ 2^k + (-1)^{k+1} & 2^k & -2^k + (-1)^k \\ 2^k + (-1)^{k+1} & 0 & (-1)^k \end{pmatrix}.$$

Au total,

$$\begin{aligned} A^k &= \begin{pmatrix} 2^k & 0 & 0 \\ 2^k + (-1)^{k+1} & 2^k & -2^k + (-1)^k \\ 2^k + (-1)^{k+1} & 0 & (-1)^k \end{pmatrix} + k \begin{pmatrix} 0 & 2^{k-1} & -2^{k-1} \\ 0 & 2^{k-1} & -2^{k-1} \\ 0 & 2^{k-1} & -2^{k-1} \end{pmatrix} \\ &= \begin{pmatrix} 2^k & k2^{k-1} & -k2^{k-1} \\ 2^k + (-1)^{k+1} & 2^k + k2^{k-1} & -2^k - k2^{k-1} + (-1)^k \\ 2^k + (-1)^{k+1} & k2^{k-1} & -k2^{k-1} + (-1)^k \end{pmatrix}. \end{aligned}$$

Si l'on connaît une réduction de Jordan de notre matrice trigonalisable $A \in \mathrm{M}_n(\mathbb{K})$, il n'est pas nécessaire de calculer la décomposition de Dunford de A pour pouvoir exprimer ses puissances successives :

Exemple 1.9.6. Reprenons la matrice

$$A = \begin{pmatrix} 3 & -1 & 1 \\ 2 & 0 & 1 \\ 1 & -1 & 2 \end{pmatrix} \in \mathrm{M}_3(\mathbb{R})$$

de l'exemple 1.8.10. Une réduction de Jordan de A est

$$P^{-1}AP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

avec $P := \begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \in \mathrm{GL}_3(\mathbb{R})$,

Ecrivons alors $P^{-1}AP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ et soit $k \in \mathbb{N}\backslash\{0\}$. Comme les matrices $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ et $\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ commutent, on a

$$\begin{aligned}
P^{-1}A^kP &= \Big(P^{-1}AP\Big)^k \\
&= \sum_{i=0}^{k} \binom{i}{k} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}^{k-i} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}^{i} \\
&= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}^{k} + k \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}^{k-1} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \text{ (la matrice } \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}^{2} \text{ est nulle)} \\
&= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2^k & 0 \\ 0 & 0 & 2^k \end{pmatrix} + k \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 2^{k-1} \\ 0 & 0 & 0 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2^k & k2^{k-1} \\ 0 & 0 & 2^k \end{pmatrix}.
\end{aligned}$$

Enfin, $P^{-1} = \begin{pmatrix} -1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & -1 & 1 \end{pmatrix}$ donc

$$A^k = P \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2^k & k2^{k-1} \\ 0 & 0 & 2^k \end{pmatrix} P^{-1} = \begin{pmatrix} 2^k + k2^{k-1} & -k2^{k-1} & k2^{k-1} \\ 2^k + k2^{k-1} - 1 & -k2^{k-1} + 1 & k2^{k-1} \\ 2^k - 1 & -2^k + 1 & 2^k \end{pmatrix}.$$

Chapitre 2

Rappels et compléments sur les espaces euclidiens

2.1 Introduction

On introduit sur les $\mathbb{R}$-espaces vectoriels de dimension finie une structure supplémentaire : le produit scalaire, notion qui généralise le produit scalaire classique sur $\mathbb{R}^2$ ou $\mathbb{R}^3$. Cette structure supplémentaire nous donne accès aux notions géométriques d'orthogonalité et de distance.

Dans ce chapitre, on abordera également la question de la réductibilité de certaines classes d'endomorphismes des espaces euclidiens. En particulier, on montre que tout endomorphisme symétrique est diagonalisable dans une base orthonormale, et que tout endomorphisme orthogonal est diagonalisable par blocs dans une base orthonormale, avec des blocs de rotations vectorielles de tailles 2 ou 1.

2.2 Produit scalaire sur un espace vectoriel réel

Soit E un espace vectoriel sur $\mathbb{R}$.

Définition 2.2.1. *Considérons une application* $\langle\cdot,\cdot\rangle : \begin{array}{ccc} E \times E & \to & \mathbb{R} \\ (v,w) & \mapsto & \langle v,w\rangle \end{array}$. *On dit que* $\langle\cdot,\cdot\rangle$ *est un produit scalaire sur E si* $\langle\cdot,\cdot\rangle$ *est*

1. *bilinéaire, i.e. pour tous* $v, w, v_1, v_2, w_1, w_2 \in E$ *et tous* $\lambda, \mu \in \mathbb{R}$, $\langle \lambda v_1 + \mu v_2, w\rangle = \lambda\langle v_1, w\rangle + \mu\langle v_2, w\rangle$ *et* $\langle v, \lambda w_1 + \mu w_2\rangle = \lambda\langle v, w_1\rangle + \mu\langle v, w_2\rangle$,
2. *symétrique, i.e. pour tous* $v, w \in E$, $\langle w, v\rangle = \langle v, w\rangle$,
3. *définie positive, i.e. pour tout* $v \in E$, $\langle v, v\rangle \geqslant 0$ *et* $\langle v, v\rangle = 0$ *si et seulement si* $v = 0_E$.

Exemple 2.2.2. 1. Soit $n \in \mathbb{N}\backslash\{0\}$, pour tous $v = (x_1, \ldots, x_n)$ et $w = (y_1, \ldots, y_n)$ dans $\mathbb{R}^n$, on définit

$$\langle v, w\rangle_{\text{can}} := x_1 y_1 + \cdots + x_n y_n.$$

L'application $\langle\cdot,\cdot\rangle_{\text{can}} : \begin{array}{ccc} \mathbb{R}^n \times \mathbb{R}^n & \to & \mathbb{R} \\ (v,w) & \mapsto & \langle v,w\rangle_{\text{can}} \end{array}$ est alors un produit scalaire sur $\mathbb{R}^n$, appelé <u>produit scalaire canonique</u> sur $\mathbb{R}^n$. Montrons le caractère défini positif de l'application $\langle\cdot,\cdot\rangle_{\text{can}}$: si $v = (x_1,\dots,x_n) \in \mathbb{R}^n$, $\langle v,v\rangle_{\text{can}} = \sum_i^n x_i^2 \geqslant 0$ et $\langle v,v\rangle_{\text{can}} = \sum_i^n x_i^2 = 0$ ssi pour tout $i \in \{1,\dots,n\}$, $x_i = 0$ ssi $v = (0,\dots,0)$.

2. Supposons que E est un espace vectoriel de dimension finie $n \in \mathbb{N}\backslash\{0\}$ et soit $\mathcal{B} = \{e_1,\dots,e_n\}$ une base de E. Pour tous vecteurs v et w de E, de coordonnées respectives $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ dans la base $\mathcal{B}$, on définit

$$\langle v,w\rangle_{\mathcal{B}} := x_1y_1 + \cdots + x_ny_n = {}^t\mathrm{Mat}_{\mathcal{B}}(v)\mathrm{Mat}_{\mathcal{B}}(w).$$

L'application $\langle\cdot,\cdot\rangle_{\mathcal{B}} : \begin{array}{ccc} E \times E & \to & \mathbb{R} \\ (v,w) & \mapsto & \langle v,w\rangle_{\mathcal{B}} \end{array}$ est alors un produit scalaire sur E, appelé <u>produit scalaire associé à la base $\mathcal{B}$</u>.

3. Soit $n \in \mathbb{N}\backslash\{0\}$. Pour toutes matrices $A = (a_{i\,j})_{1\leqslant i,j\leqslant n}$ et $B = (b_{i\,j})_{1\leqslant i,j\leqslant n}$ dans $\mathrm{M}_n(\mathbb{R})$, on définit

$$\langle A,B\rangle := \sum_{1\leqslant i,j\leqslant n} a_{i\,j}b_{i\,j} = \mathrm{Tr}\left({}^tA\,B\right).$$

L'application $\langle\cdot,\cdot\rangle : \begin{array}{ccc} \mathrm{M}_n(\mathbb{R}) \times \mathrm{M}_n(\mathbb{R}) & \to & \mathbb{R} \\ (A,B) & \mapsto & \langle A,B\rangle \end{array}$ est alors un produit scalaire sur $\mathrm{M}_n(\mathbb{R})$: il s'agit du produit scalaire associé à la base canonique $\{E_{i\,j}\}_{1\leqslant i,j\leqslant n}$ de $\mathrm{M}_n(\mathbb{R})$.

4. Soit $n \in \mathbb{N}\backslash\{0\}$. Pour tous polynômes P et Q de $\mathbb{R}_n[X]$, on définit

$$\langle P,Q\rangle := \int_0^1 P(t)Q(t)\mathrm{d}t.$$

L'application $\langle\cdot,\cdot\rangle : \begin{array}{ccc} \mathbb{R}_n[X] \times \mathbb{R}_n[X] & \to & \mathbb{R} \\ (P,Q) & \mapsto & \langle P,Q\rangle \end{array}$ est alors un produit scalaire sur $\mathbb{R}_n[X]$. La bilinéarité de l'application $\langle\cdot,\cdot\rangle$ provient de la linéarité de l'intégrale. Montrons que $\langle\cdot,\cdot\rangle$ est bien définie positive. Soit $P \in \mathbb{R}_n[X]$, on a $\langle P,P\rangle = \int_0^1 (P(t))^2\mathrm{d}t \geqslant 0$ (l'intégrale sur un segment d'une fonction positive est positive). Si $\langle P,P\rangle = \int_0^1 (P(t))^2\mathrm{d}t = 0$, comme la fonction $\mathbb{R} \to \mathbb{R}$; $t \mapsto P(t)^2$ est continue et positive, on a, pour tout $t \in [0,1]$, $(P(t))^2 = 0$ (l'intégrale d'une fonction continue et positive sur un segment est nulle si et seulement si la fonction est identiquement nulle sur ce segment) et donc, pour tout $t \in [0,1]$, $P(t) = 0$, donc le polynôme P est nul (un polynôme ayant une infinité de racines est nécessairement le polynôme nul).

Remarque 2.2.3. Si $\langle \cdot, \cdot \rangle$ est un produit scalaire sur E et si F est un sous-espace vectoriel de E, la restriction

$$\langle \cdot, \cdot \rangle_{|F\times F} : \begin{array}{ccc} F \times F & \to & \mathbb{R} \\ (v, w) & \mapsto & \langle v, w \rangle \end{array}$$

est un produit scalaire sur F.

Définition 2.2.4. *Si E est un espace vectoriel de dimension finie muni d'un produit scalaire $\langle \cdot, \cdot \rangle$, le couple $(E, \langle \cdot, \cdot \rangle)$ est appelé espace euclidien.*

Remarque 2.2.5. D'après la remarque 2.2.3, si $(E, \langle\ ,\ \rangle)$ est un espace euclidien et si F est un sous-espace vectoriel de E, alors $\left(F, \langle\ ,\ \rangle_{|F\times F}\right)$ est également un espace euclidien. On le notera simplement $(F, \langle\ ,\ \rangle)$.

On suppose dans la suite de cette section que E est un espace vectoriel de dimension finie muni d'un produit scalaire $\langle \cdot, \cdot \rangle$. Pour tout vecteur v de E, $\langle v, v \rangle \geqslant 0$ et on note alors $\|v\| := \sqrt{\langle v, v \rangle}$. Une première propriété importante des espaces euclidiens est l'inégalité de Cauchy-Schwarz présentée ci-dessous. Cette inégalité permet en particulier de montrer que l'application qui à tout vecteur v de E associe $\|v\| \in [0, +\infty[$ est une norme.

Lemme 2.2.6 (Inégalité de Cauchy-Schwarz)**.** *Pour tous vecteurs v et w de E, on a l'inégalité*

$$|\langle v, w \rangle| \leqslant \|v\| \|w\|,$$

et $|\langle v, w \rangle| = \|v\| \|w\|$ si et seulement si les vecteurs v et w sont liés.

Démonstration. Soient $v, w \in E$.

Si w est le vecteur nul 0_E de E, on a

$$\langle v, w \rangle = \langle v, 0_E \rangle = \langle v, 0 \cdot 0_E \rangle = 0 \cdot \langle v, 0_E \rangle = 0$$

et $\|w\| = \sqrt{\langle w, w \rangle} = \sqrt{\langle 0_E, 0_E \rangle} = 0$. L'inégalité ci-dessus est donc vérifiée : il s'agit d'une égalité et on a $w = 0_E = 0 \cdot v$.

On suppose maintenant que $w \neq 0_E$. Soit alors $\lambda \in \mathbb{R}$, on a $\|v + \lambda w\|^2 \geqslant 0$. Or

$$\begin{aligned} \|v + \lambda w\|^2 &= \langle v + \lambda w, v + \lambda w \rangle \\ &= \langle v, v \rangle + \lambda \langle v, w \rangle + \lambda \langle w, v \rangle + \lambda^2 \langle w, w \rangle \\ &= \|v\|^2 + 2\lambda \langle v, w \rangle + \lambda^2 \|w\|^2. \end{aligned}$$

Ainsi, pour tout $\lambda \in \mathbb{R}$, $\|w\|^2 \lambda^2 + 2\langle v, w \rangle \lambda + \|v\|^2 \geqslant 0$, en d'autres termes, la fonction polynomiale de degré deux

$$\begin{array}{ccc} \mathbb{R} & \to & \mathbb{R} \\ \lambda & \mapsto & \|w\|^2 \lambda^2 + 2\langle v, w \rangle \lambda + \|v\|^2 \end{array}$$

(remarquons que $\|w\|^2 \neq 0$ car $\langle w, w \rangle \neq 0$ car $w \neq 0_E$) est positive sur tout $\mathbb{R}$, ce qui est équivalent au fait que le discriminant associé $4\langle v, w \rangle^2 - 4\|w\|^2 \|v\|^2$ soit négatif ou nul. Ainsi, on a $\langle v, w\rangle^2 \leqslant \|v\|^2\|w\|^2$ i.e. $|\langle v, w\rangle| \leqslant \|v\|\|w\|$.

De plus, si $|\langle v, w\rangle| = \|v\|\|w\| \Leftrightarrow \langle v, w\rangle^2 = \|v\|^2\|w\|^2$, le discriminant associé à la fonction polynomiale du second degré ci-dessus est nul : le polynôme associé possède donc une racine (double) $\lambda_0 \in \mathbb{R}$. On a ainsi

$$\langle v + \lambda_0 w, v + \lambda_0 w\rangle = \|v + \lambda_0 w\|^2 = \|w\|^2\lambda_0^2 + 2\langle v, w\rangle\lambda_0 + \|v\|^2 = 0,$$

et donc $v + \lambda_0 w = 0_E$. En particulier, les vecteurs v et w sont liés.

Réciproquement, supposons qu'il existe $(\mu_1, \mu_2) \in \mathbb{R}^2 \backslash \{(0, 0)\}$ tel que $\mu_1 v + \mu_2 w = 0$. Si $\mu_2 = 0$, nécessairement $v = 0_E$ et, comme ci-dessus, $|\langle v, w\rangle| = 0 = \|v\|\|w\|$. Si $\mu_2 \neq 0$, on a $w = \frac{\mu_1}{\mu_2} v$ et alors

$$\begin{aligned}
\langle v, w\rangle^2 &= \left\langle v, \frac{\mu_1}{\mu_2} v\right\rangle^2 \\
&= \left(\frac{\mu_1}{\mu_2}\right)^2 \langle v, v\rangle^2 \\
&= \langle v, v\rangle \left\langle \frac{\mu_1}{\mu_2} v, \frac{\mu_1}{\mu_2} v\right\rangle \\
&= \|v\|^2\|w\|^2.
\end{aligned}$$

□

Corollaire et Définition 2.2.7. *L'application* $\|\cdot\| : \begin{array}{ccc} E & \to & [0, +\infty[\\ v & \mapsto & \|v\| \end{array}$ *est une norme, i.e.*

1. *pour tous $v \in E$ et $\lambda \in \mathbb{R}$, $\|\lambda v\| = |\lambda|\|v\|$,*
2. *pour tout $v \in E$, $\|v\| = 0$ si et seulement si $v = 0_E$,*
3. *pour tous $v, w \in E$, $\|v + w\| \leqslant \|v\| + \|w\|$ (inégalité de Minkowski).*

En conséquence, le couple $(E, \|\cdot\|)$ est un espace vectoriel normé. La norme $\|\cdot\|$ est appelée <u>*norme euclidienne*</u> *associée au produit scalaire $\langle\cdot, \cdot\rangle$.*

Démonstration. 1. Soient $v \in E$ et $\lambda \in \mathbb{R}$, alors

$$\|\lambda v\| = \sqrt{\langle \lambda v, \lambda v\rangle} = \sqrt{\lambda^2\langle v, v\rangle} = \sqrt{\lambda^2}\sqrt{\langle v, v\rangle} = |\lambda|\|v\|.$$

2. Soit $v \in E$, alors $\|v\| = \sqrt{\langle v, v\rangle} = 0$ ssi $\langle v, v\rangle = 0$ ssi $v = 0_E$.

3. Soient $v, w \in E$. Alors

$$\begin{aligned}
\|v + w\|^2 &= \langle v + w, v + w\rangle \\
&= \langle v, v\rangle + \langle v, w\rangle + \langle w, v\rangle + \langle w, w\rangle \\
&= \|v\|^2 + 2\langle v, w\rangle + \|w\|^2 \\
&\leqslant \|v\|^2 + 2|\langle v, w\rangle| + \|w\|^2 \\
&\leqslant \|v\|^2 + 2\|v\|\|w\| + \|w\|^2 \text{ (par l'inégalité de Cauchy-Schwarz)} \\
&= (\|v\| + \|w\|)^2
\end{aligned}$$

et donc $\|v+w\| \leqslant \|v\|+\|w\|$.

□

Remarque 2.2.8. • D'après les premières égalités ci-dessus, on a, pour tous $v, w \in E$,

$$\langle v, w\rangle = \frac{1}{2}\left(\|v+w\|^2 - \|v\|^2 - \|w\|^2\right).$$

- Soient $v, w \in E$, alors $\|v+w\| = \|v\|+\|w\|$ si et seulement si $w = 0_E$ ou $v = \mu\, w$ avec $\mu \in [0; +\infty[$.

  En effet, supposons que $\|v+w\| = \|v\|+\|w\|$ et que $w \neq 0_E$, alors

$$\begin{aligned} \|v+w\|^2 = (\|v\|+\|w\|)^2 \quad &\text{i.e.} \quad \|v\|^2 + 2\langle v, w\rangle + \|w\|^2 = \|v\|^2 + 2\|v\|\|w\| + \|w\|^2 \\ &\text{i.e.} \quad \langle v, w\rangle = \|v\|\|w\| \end{aligned}$$

  et donc, d'après la démonstration du cas d'égalité de l'inégalité de Cauchy-Schwarz (cf. preuve du lemme 2.2.6), $v = -\lambda_0 w$ où λ_0 est la racine double du polynôme de degré deux de discriminant nul $\|w\|^2 X^2 + 2\langle v, w\rangle X + \|v\|^2$, i.e. $\lambda_0 = -\frac{2\langle v,w\rangle}{2\|w\|^2} = -\frac{\langle v,w\rangle}{\|w\|^2} = -\frac{\|v\|\|w\|}{\|w\|^2}$. En particulier, $-\lambda_0 \geqslant 0$.

  Réciproquement, si $w = 0_E$, alors $\|v+w\| = \|v\| = \|v\|+\|w\|$, et si $v = \mu\, w$ avec $\mu \in [0; +\infty[$, alors

$$\begin{aligned} \|v+w\| &= \|\mu w + w\| \\ &= \|(\mu+1)w\| \\ &= (\mu+1)\|w\| \ (|\mu+1| = \mu+1 \text{ car } \mu \geqslant 0) \\ &= \mu\|w\| + \|w\| \\ &= \|\mu w\| + \|w\| \ (\mu = |\mu| \text{ car } \mu \geqslant 0) \\ &= \|v\| + \|w\|. \end{aligned}$$

2.3 Orthogonalité dans les espaces euclidiens, bases orthogonales, bases orthonormales

Soit $(E, \langle\cdot,\cdot\rangle)$ un espace euclidien.

La structure supplémentaire qu'apporte le produit scalaire $\langle\cdot,\cdot\rangle$ à l'espace vectoriel de dimension finie E nous permet d'y introduire une notion d'orthogonalité :

Définition 2.3.1. *Soient v et w deux vecteurs de E. On dit que v et w sont orthogonaux si $\langle v, w\rangle = 0$. Dans ce cas, on dira aussi que v est orthogonal à w et que w est orthogonal à v.*

Exemple 2.3.2. Dans $\mathbb{R}^3$ muni du produit scalaire canonique, les vecteurs $(1, -1, 2)$ et $(1, 3, 1)$ sont orthogonaux.

Remarquons que le théorème de Pythagore peut être étendu à tout espace euclidien :

Lemme 2.3.3 (Théorème de Pythagore). *Soient v et w deux vecteurs de E. Alors v et w sont orthogonaux si et seulement si $\|v+w\|^2 = \|v\|^2 + \|w\|^2$.*

Démonstration. On a $\|v+w\|^2 = \|v\|^2 + 2\langle v, w\rangle + \|w\|^2$ et donc $\|v+w\|^2 = \|v\|^2 + \|w\|^2$ ssi $\langle v, w\rangle = 0$ ssi v et w sont orthogonaux. □

Notons $n := \dim(E)$ et soit maintenant $\{v_1, \ldots, v_n\}$ une base de E. Si v et w sont deux vecteurs de E de coordonnées respectives $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ dans cette base, alors

$$\langle v, w\rangle = \left\langle \sum_{i=1}^n x_i v_i, \sum_{j=1}^n y_j v_j \right\rangle = \sum_{1 \leqslant i,j \leqslant n} x_i y_j \langle v_i, v_j\rangle.$$

On aimerait pouvoir considérer une base de E dans laquelle cette expression du produit scalaire $\langle \cdot, \cdot \rangle$ sur E soit la plus simple possible. Cela nous mène à la notion de base orthogonale et de base orthonormale :

Définition 2.3.4. *Soit $\mathcal{B} = \{e_1, \ldots, e_n\}$ une base de E.*

1. *On dit que $\mathcal{B}$ est une <u>base orthogonale de E</u> si pour tous $i, j \in \{1, \ldots, n\}$ tels que $i \neq j$, on a $\langle e_i, e_j\rangle = 0$.*

2. *On dit que $\mathcal{B}$ est une <u>base orthonormale de E</u> si $\mathcal{B}$ est une base orthogonale de E et si tous les vecteurs de $\mathcal{B}$ sont de norme euclidienne 1. Autrement dit, $\mathcal{B}$ est une base orthonormale de E si et seulement si pour tous $i, j \in \{1, \ldots, n\}$, $\langle e_i, e_j\rangle = \delta_{i\,j}$ (pour tout $v \in E$, $\|v\| = 1$ ssi $\langle v, v\rangle = 1$).*

Exemple 2.3.5.

- La base canonique de $\mathbb{R}^n$ est une base orthonormale pour le produit scalaire canonique sur $\mathbb{R}^n$ (exemple 2.2.2 1.).
- Par définition, une base $\mathcal{B}$ de E est une base orthonormale pour le produit scalaire associé $\langle \cdot, \cdot \rangle_{\mathcal{B}}$ (exemple 2.2.2 2.).

Remarque 2.3.6.

1. Si $\mathcal{B} = \{e_1, \ldots, e_n\}$ est une base orthonormale pour $\langle \cdot, \cdot \rangle$ et si v et w sont deux vecteurs de E de coordonnées respectives $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ dans $\mathcal{B}$, alors

$$\langle v, w\rangle = \sum_{1 \leqslant i,j \leqslant n} x_i y_j \langle e_i, e_j\rangle = \sum_{i=1}^n x_i y_i.$$

En particulier, pour tout $i \in \{1, \ldots, n\}$, $\langle v, e_i\rangle = x_i$ et donc

$$v = \sum_{i=1}^n \langle v, e_i\rangle\, e_i.$$

2. Si la famille $\{\epsilon_1, \ldots, \epsilon_n\}$ est une base orthogonale de E, alors la famille $\{e_1, \ldots, e_n\}$ avec, pour tout $i \in \{1, \ldots, n\}$, $e_i := \frac{\epsilon_i}{\|\epsilon_i\|}$, est une base orthonormale de E. En effet, pour tout $i \in \{1, \ldots, n\}$, $\|e_i\| = \left\| \frac{\epsilon_i}{\|\epsilon_i\|} \right\| = \frac{1}{\|\epsilon_i\|}\|\epsilon_i\| = 1$.

3. Plus généralement, on appellera famille orthogonale toute famille finie de vecteurs non nuls de E deux à deux orthogonaux, et famille orthonormale toute famille orthogonale $\{v_1, \ldots, v_p\}$ de E telle que, pour tout $i \in \{1, \ldots, p\}$, $\|v_i\| = 1$. Il est à remarquer que toute famille orthogonale de E est libre : soit $\{w_1, \ldots, w_m\}$ une famille orthogonale de E et soient $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$ tels que $\lambda_1 w_1 + \cdots + \lambda_m w_m = 0_E$, alors, pour tout $i \in \{1, \ldots, m\}$,
$$0 = \left\langle w_i, \sum_{j=1}^{m} \lambda_j w_j \right\rangle = \sum_{j=1}^{m} \lambda_j \langle w_i, w_j\rangle = \lambda_i \langle w_i, w_i\rangle$$
et donc, comme $\langle w_i, w_i\rangle \neq 0$ (car $w_i \neq 0_E$), $\lambda_i = 0$. En particulier, étant libre, toute famille orthogonale possède au plus $n = \dim(E)$ éléments.

Un résultat essentiel de la théorie des espaces euclidiens est qu'il est toujours possible de construire, de façon algorithmique, une base orthonormale pour l'espace euclidien considéré :

Théorème 2.3.7 (Procédé d'orthonormalisation de Gram-Schmidt)**.** *Soit $\{v_1, \ldots, v_p\}$ une famille libre de E. On peut construire, de façon algorithmique, une famille orthonormale $\{e_1, \ldots, e_p\}$ de E telle que, pour tout $k \in \{1, \ldots, p\}$,* $\mathrm{Vect}\{e_1, \ldots, e_k\} = \mathrm{Vect}\{v_1, \ldots, v_k\}$ *(en particulier, $\{e_1, \ldots, e_p\}$ est donc une base orthonormale de* $\mathrm{Vect}\{v_1, \ldots, v_p\}$*).*

Démonstration. On construit, de façon récursive, une base orthogonale $\{\epsilon_1, \ldots, \epsilon_p\}$ pour $\mathrm{Vect}\{v_1, \ldots, v_p\}$. On en déduit immédiatement une base orthonormale en "normalisant" les vecteurs obtenus (i.e. en les multipliant chacun par l'inverse de leur norme comme dans la remarque 2.3.6 2.).

Le procédé récursif est le suivant : pour $1 \leqslant k \leqslant p-1$, on suppose que l'on a déjà construit des vecteurs $\epsilon_1, \ldots, \epsilon_k \in E$ orthogonaux deux à deux tels que $\mathrm{Vect}\{\epsilon_1, \ldots, \epsilon_k\} = \mathrm{Vect}\{v_1, \ldots, v_k\}$. On recherche alors un vecteur ϵ_{k+1} de E de la forme
$$\epsilon_{k+1} = v_{k+1} + \lambda_1\epsilon_1 + \cdots + \lambda_k\epsilon_k \in \mathrm{Vect}\{\epsilon_1, \ldots, \epsilon_k, v_{k+1}\} = \mathrm{Vect}\{v_1, \ldots, v_k, v_{k+1}\}$$
tel que, pour tout $i \in \{1, \ldots, k\}$, $\langle \epsilon_{k+1}, \epsilon_i\rangle = 0$.

Or, pour $i \in \{1, \ldots, k\}$,
$$\begin{aligned}\langle \epsilon_{k+1}, \epsilon_i\rangle = 0 \quad &\Leftrightarrow \quad \langle v_{k+1} + \lambda_1\epsilon_1 + \cdots + \lambda_k\epsilon_k, \epsilon_i\rangle = 0\\ &\Leftrightarrow \quad \langle v_{k+1}, \epsilon_i\rangle + \lambda_i\langle \epsilon_i, \epsilon_i\rangle = 0\\ &\Leftrightarrow \quad \lambda_i = -\frac{\langle v_{k+1}, \epsilon_i\rangle}{\|\epsilon_i\|^2}\end{aligned}$$
Ainsi, par construction, la famille $\{\epsilon_1, \ldots, \epsilon_{k+1}\}$ avec $\epsilon_{k+1} := v_{k+1} - \sum_{i=1}^{k} \frac{\langle v_{k+1}, \epsilon_i\rangle}{\|\epsilon_i\|^2}\epsilon_i$ est orthogonale et engendre bien $\mathrm{Vect}\{v_1, \ldots, v_k, v_{k+1}\}$ car $\mathrm{Vect}\{v_1, \ldots, v_k, v_{k+1}\} = \mathrm{Vect}\{\epsilon_1, \ldots, \epsilon_k, v_{k+1}\} = \mathrm{Vect}\{\epsilon_1, \ldots, \epsilon_k, \epsilon_{k+1}\}$ (car $\epsilon_{k+1} = v_{k+1} - \sum_{i=1}^{k} \frac{\langle v_{k+1}, \epsilon_i\rangle}{\|\epsilon_i\|^2}\epsilon_i \in \mathrm{Vect}\{\epsilon_1, \ldots, \epsilon_k, v_{k+1}\}$ et $v_{k+1} = \epsilon_{k+1} + \sum_{i=1}^{k} \frac{\langle v_{k+1}, \epsilon_i\rangle}{\|\epsilon_i\|^2}\epsilon_i \in \mathrm{Vect}\{\epsilon_1, \ldots, \epsilon_k, \epsilon_{k+1}\}$). □

Remarque 2.3.8. Au terme de l'étape $k+1$ du procédé ci-dessus, on peut remplacer ϵ_{k+1} par n'importe quel vecteur non nul ϵ'_{k+1} de la droite vectorielle engendrée par ϵ_{k+1} : la famille $\{\epsilon_1, \ldots, \epsilon_k, \epsilon'_{k+1}\}$ ainsi obtenue reste orthogonale et engendre toujours Vect $\{v_1, \ldots, v_{k+1}\}$. Cela peut être utile pour simplifier les calculs (notamment pour éviter de manipuler des fractions).

Exemple 2.3.9. On détermine une base orthonormale pour le sous-espace vectoriel F de $\mathbb{R}^4$ engendré par les vecteurs $v_1 = (1,1,0,0)$, $v_2 = (1,0,-1,1)$ et $v_3 = (0,1,1,1)$. La famille $\{v_1, v_2, v_3\}$ est libre (c'est donc une base de $F = \mathrm{Vect}\{v_1, v_2, v_3\}$) et on applique le procédé d'orthonormalisation de Gram-Schmidt pour déterminer une base orthonormale de F.

On pose

$$\begin{array}{rcl}
\epsilon_1 & := & v_1 = (1,1,0,0) \\
\epsilon_2 & := & v_2 - \dfrac{\langle v_2, \epsilon_1 \rangle}{\|\epsilon_1\|^2}\epsilon_1 = v_2 - \dfrac{1}{2}\epsilon_1 = \left(\dfrac{1}{2}, -\dfrac{1}{2}, -1, 1\right) = \dfrac{1}{2}(1,-1,-2,2) \\
\epsilon'_2 & := & (1,-1,-2,2) \text{ (pour simplifier les calculs)} \\
\epsilon_3 & := & v_3 - \dfrac{\langle v_3, \epsilon_1 \rangle}{\|\epsilon_1\|^2}\epsilon_1 - \dfrac{\langle v_3, \epsilon'_2 \rangle}{\|\epsilon'_2\|^2}\epsilon'_2 = v_3 - \dfrac{1}{2}\epsilon_1 + \dfrac{1}{10}\epsilon'_2 = \left(-\dfrac{2}{5}, \dfrac{2}{5}, \dfrac{4}{5}, \dfrac{6}{5}\right) = \dfrac{2}{5}(-1,1,2,3) \\
\epsilon'_3 & := & (-1,1,2,3)
\end{array}$$

et la famille $\{\epsilon_1, \epsilon'_2, \epsilon'_3\}$ est alors une base orthogonale de F. La famille

$$\left\{\frac{1}{\sqrt{2}}(1,1,0,0), \frac{1}{\sqrt{10}}(1,-1,-2,2), \frac{1}{\sqrt{15}}(-1,1,2,3)\right\}$$

ensuite obtenue par normalisation des vecteurs $\{\epsilon_1, \epsilon'_2, \epsilon'_3\}$ est une base orthonormale de F.

Corollaire 2.3.10. *Il existe une base orthonormale pour l'espace euclidien* $(E, \langle \cdot, \cdot \rangle)$

Démonstration. Soit $\{v_1, \ldots, v_n\}$ une base de E. En particulier, la famille $\{v_1, \ldots, v_n\}$ est libre et, d'après le théorème 2.3.7, on peut construire une base orthonormale pour $\mathrm{Vect}\{v_1, \ldots, v_n\} = E$. □

Remarque 2.3.11. Soit $\mathcal{B} = \{e_1, \ldots, e_n\}$ une base orthonormale pour $\langle \cdot, \cdot \rangle$. Alors $\langle \cdot, \cdot \rangle = \langle \cdot, \cdot \rangle_{\mathcal{B}}$ (exemple 2.2.2 2.). En effet, pour tous vecteurs v et w de E, de coordonnées respectives $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ dans la base $\mathcal{B}$, on a $\langle v, w \rangle = \displaystyle\sum_{i=1}^{n} x_i y_i = \langle v, w \rangle_{\mathcal{B}}$.

Le choix d'une base orthonormale pour l'espace euclidien $(E, \langle \cdot, \cdot \rangle)$ permet d'identifier E avec l'espace euclidien $(\mathbb{R}^n, \langle \cdot, \cdot \rangle_{\mathrm{can}})$ muni du produit scalaire canonique. Précisément :

Corollaire 2.3.12. *Il existe un isomorphisme (non canonique)* $\psi : E \to \mathbb{R}^n$ *tel que, pour tous vecteurs* v *et* w *de* E, $\langle \psi(v), \psi(w) \rangle_{\mathrm{can}} = \langle v, w \rangle$.

Démonstration. Soit $\mathcal{B} = \{e_1, \ldots, e_n\}$ une base orthonormale pour $\langle \cdot, \cdot \rangle$ et notons $\mathcal{B}' = \{e'_1, \ldots, e'_n\}$ la base canonique de $\mathbb{R}^n$. Notons ensuite $\psi_{\mathcal{B}}$ l'application linéaire de E dans $\mathbb{R}^n$ qui, pour tout $i \in \{1, \ldots, n\}$, associe e'_i à e_i. Autrement dit, $\psi_{\mathcal{B}}$ est l'application qui à tout vecteur v de E de coordonnées $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ dans la base $\mathcal{B}$ associe le vecteur $(x_1, \ldots, x_n)$ de $\mathbb{R}^n$. Il s'agit d'un isomorphisme linéaire et, si v et w sont deux vecteurs de E de coordonnées respectives $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ dans $\mathcal{B}$, on a

$$\langle \psi_{\mathcal{B}}(v), \psi_{\mathcal{B}}(w) \rangle_{\text{can}} = \langle (x_1, \ldots, x_n), (y_1, \ldots, y_n) \rangle_{\text{can}} = \sum_{i=1}^{n} x_i y_i = \langle v, w \rangle$$

(car $\mathcal{B}$ est une base orthonormale pour $\langle \cdot, \cdot \rangle$). □

Remarque 2.3.13. Le théorème 2.3.7 permet également de montrer que toute famille orthonormale de E peut être complétée en une base orthonormale de E. En effet, soit $\{e_1, \ldots, e_p\}$ une famille orthonormale de E, on la complète en une base $\{e_1, \ldots, e_p, v_{p+1}, \ldots, v_n\}$ de E à laquelle on applique le procédé d'orthonormalisation de Gram-Schmidt : comme les vecteurs $e_1, \ldots, e_p$ sont déjà orthogonaux deux à deux et tous de norme 1, on est amené à construire des vecteurs $e_{p+1}, \ldots, e_n$ de E tels que la famille $\{e_1, \ldots, e_p, e_{p+1}, \ldots, e_n\}$ soit une base orthonormale de E.

2.4 Orthogonal d'un sous-ensemble, orthogonal d'un sous-espace vectoriel

Soit $(E, \langle \cdot, \cdot \rangle)$ un espace euclidien et soit A un sous-ensemble non vide de E. On considère l'ensemble des vecteurs de E orthogonaux à tous les vecteurs contenus dans A :

Définition 2.4.1. *On note $A^\perp$ l'ensemble*

$$\{v \in E \mid \textit{pour tout } w \in A, \langle v, w \rangle = 0\},$$

que l'on appelle l'orthogonal de A (par rapport à $\langle \cdot, \cdot \rangle$).

Exemple 2.4.2. Dans $\mathbb{R}^3$ muni du produit scalaire canonique, l'orthogonal du sous-ensemble de $\mathbb{R}^3$ réduit au vecteur $(-2, 5, 3)$ est

$$\{(-2,5,3)\}^\perp = \left\{(x,y,z) \in \mathbb{R}^3 \mid \langle (x,y,z), (-2,5,3) \rangle_{\text{can}} = -2x + 5y + 3z = 0\right\}.$$

Dans l'exemple ci-dessus, on peut remarquer que $\{(-2, 5, 3)\}^\perp$ est un sous-espace vectoriel de $\mathbb{R}^3$. L'orthogonal d'un sous-ensemble d'un espace euclidien en est en fait toujours un sous-espace vectoriel :

Lemme 2.4.3. *Soit A un sous-ensemble non vide de E. Alors $A^\perp$ est un sous-espace vectoriel de E (même si A ne l'est pas !).*

Démonstration. $0_E \in A^\perp$ car, pour tout $w \in A$, $\langle 0_E, w\rangle = 0$, et, si $v_1, v_2 \in A^\perp$ et $\lambda, \mu \in \mathbb{R}$, on a, pour tout $w \in A$,

$$\langle \lambda v_1 + \mu v_2, w\rangle = \lambda\langle v_1, w\rangle + \mu\langle v_2, w\rangle = 0.$$

Ainsi, $A^\perp$ est bien un sous-espace vectoriel de E. □

Remarque 2.4.4.
- Si A est un sous-ensemble de E, on a $A^\perp = (\mathrm{Vect}(A))^\perp$.
- On a $\{0_E\}^\perp = E$ (car, pour tout $v \in E$, $\langle v, 0_E\rangle = 0$) et $E^\perp = \{0_E\}$ (en effet, si $v \in E^\perp$, $\langle v, v\rangle = 0$, car $v \in E^\perp$ et $v \in E$, et donc $v = 0_E$).

Si maintenant F est un sous-espace vectoriel de E, on va montrer que l'on peut décomposer E en la somme directe de F et de son orthogonal $F^\perp$:

Proposition 2.4.5. *Soit F un sous-espace vectoriel de E. On a :*

1. $E = F \oplus F^\perp$ *(en particulier,* $\dim\left(F^\perp\right) = \dim(E) - \dim(F)$*),*
2. $\left(F^\perp\right)^\perp = F$.

Démonstration.
1. On note $p := \dim(F)$. Soit $\{e_1, \ldots, e_p\}$ une base orthonormale de F que l'on complète en une base orthonormale $\{e_1, \ldots, e_n\}$ de E. On a alors $E = F \oplus \mathrm{Vect}\{e_{p+1}, \ldots, e_n\}$. Montrons que $\mathrm{Vect}\{e_{p+1}, \ldots, e_n\} = F^\perp$.

   Soit $v \in E$. Commençons par remarquer que $v \in F^\perp$ ssi pour tout $i \in \{1, \ldots, p\}$, $\langle v, e_i\rangle = 0$ (car $\{e_1, \ldots, e_p\}$ est une famille génératrice de F). De plus, comme $\{e_1, \ldots, e_n\}$ est une base orthonormale de E, $v = \sum_{i=1}^{n}\langle v, e_i\rangle e_i$ (remarque 2.3.6 1.) et, ainsi,

$$\begin{aligned} v \in F^\perp \quad & \text{ssi} \quad \forall i \in \{1, \ldots, p\}, \langle v, e_i\rangle = 0 \\ & \text{ssi} \quad v = \sum_{i=p+1}^{n} \langle v, e_i\rangle e_i \ (\{e_1, \ldots, e_n\} \text{ est une base de } E) \\ & \text{ssi} \quad v \in \mathrm{Vect}\{e_{p+1}, \ldots, e_n\} \ (\{e_{p+1}, \ldots, e_n\} \text{ est une base orthonormale de } \mathrm{Vect}\{e_{p+1}, \ldots, e_n\}). \end{aligned}$$

2. On a $F \subset \left(F^\perp\right)^\perp$ car, si $v \in F$, pour tout $w \in F^\perp$, $\langle v, w\rangle = \langle w, v\rangle = 0$. De plus,

$$\dim\left(\left(F^\perp\right)^\perp\right) = \dim(E) - \dim\left(F^\perp\right) = \dim(E) - (\dim(E) - \dim(F)) = \dim(F).$$

   Par conséquent, $F = \left(F^\perp\right)^\perp$.

□

Soit F un sous-espace vectoriel de E. Comme E se décompose en la somme directe de F et de son orthogonal $F^\perp$, on peut considérer la projection de E sur F parallèlement à $F^\perp$:

Définition 2.4.6. *On appelle projection orthogonale sur F la projection de E sur F parallèlement à $F^\perp$. Il s'agit de l'application linéaire surjective qui à tout vecteur $v = w + u$ de E avec $w \in F$ et $u \in F^\perp$ associe sa composante w dans F. On note p_F cette application.*

Remarque 2.4.7. • Si $\{e_1, \dots, e_p\}$ est une base de F et $\{e_{p+1}, \dots, e_n\}$ est une base de $F^\perp$ alors $\mathcal{B} := \{e_1, \dots, e_p, e_{p+1}, \dots, e_n\}$ est une base de E dans laquelle la matrice représentative de p_F est

$$\mathrm{Mat}_{\mathcal{B}}(p_F) = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} = \begin{pmatrix} I_p & 0_{p,n-p} \\ 0_{n-p,p} & 0_{n-p,n-p} \end{pmatrix}.$$

• Si $\{e_1, \dots, e_p\}$ est une base orthonormale de F alors, pour tout $v \in E$,

$$p_F(v) = \sum_{i=1}^{p} \langle v, e_i \rangle e_i.$$

• Soient $v \in E$ et $w \in F$. Alors $w = p_F(v)$ si et seulement si $v - w \in F^\perp$: en effet, on a $v = v - w + w$ avec $w \in F$ et $E = F^\perp \oplus F$.

La projection orthogonale permet notamment de calculer explicitement la distance euclidienne d'un vecteur de E à un sous-espace vectoriel de E. On donne tout d'abord la définition de cette notion de distance euclidienne :

Définition 2.4.8. *Si v et w sont deux vecteurs de E, on définit la distance euclidienne de v à w comme étant le réel positif ou nul $d(v, w) := \|w - v\|$.*

Si v est un vecteur de E et A un sous-ensemble non vide de E, on définit également la distance euclidienne de v à A comme étant le réel positif ou nul $d(v, A) := \inf_{w \in A} d(v, w) = \inf_{w \in A} \|w - v\|$.

Remarque 2.4.9. • La distance euclidienne d'un vecteur de E à un autre est bien une distance (il s'agit de la distance induite par la norme $|| \cdot ||$).

• Pour v un vecteur de E et A un sous-ensemble non vide de E, la borne inférieure de l'ensemble $\{d(v, w) \mid w \in A\}$ existe bien et est positive ou nulle : ce sous-ensemble de $\mathbb{R}$ est non vide et minorée par 0.

Proposition 2.4.10. *Soit $v \in E$. On a*

$$d(v, F) = \|v - p_F(v)\|.$$

Démonstration. Soit $w \in F$. On a

$$\|v-w\|^2 = \|v-p_F(v)+p_F(v)-w\|^2 = \|v-p_F(v)\|^2 + \|p_F(v)-w\|^2$$

par le théorème de Pythagore (lemme 2.3.3) car $v - p_F(v) \in F^\perp$ (par définition de la projection orthogonale : il existe un unique vecteur $u \in F^\perp$ tel que $v = p_F(v)+u$) et $p_F(v) - w \in F$ (car F est un sous-espace vectoriel de E).

Ainsi, pour tout $w \in F$, $\|v-w\|^2 \geqslant \|v-p_F(v)\|^2$ donc $\|v-w\| \geqslant \|v-p_F(v)\|$ donc $\inf\{\|v-w\| \mid w \in F\} \geqslant \|v-p_F(v)\|$. De plus, comme $p_F(v) \in F$, $\inf\{\|v-w\| \mid w \in F\} \leqslant \|v-p_F(v)\|$. Au total, $\|v-p_F(v)\| = \inf\{\|v-w\| \mid w \in F\} = d(v,F)$. □

Remarque 2.4.11. 1. En particulier, la borne inférieure de l'ensemble $\{d(v,w) \mid w \in F\}$ est un minimum.

2. Si $v \in E$ et $w \in F$, $d(v,F) = \|v-w\|$ si et seulement si $w = p_F(v)$: on a

$$\begin{aligned} d(v,F) = \|v-w\| \quad & \text{ssi} \quad \|v-w\| = \|v-p_F(v)\| \\ & \text{ssi} \quad \|v-w\|^2 = \|v-p_F(v)\|^2 \\ & \text{ssi} \quad \|v-p_F(v)\|^2 + \|p_F(v)-w\|^2 = \|v-p_F(v)\|^2 \\ & \text{ssi} \quad \|p_F(v)-w\|^2 \\ & \text{ssi} \quad p_F(v)-w = 0_E \\ & \text{ssi} \quad w = p_F(v). \end{aligned}$$

On définit également la symétrie orthogonale par rapport à F :

Définition 2.4.12. *On appelle symétrie orthogonale par rapport à F la symétrie par rapport à F parallèlement à $F^\perp$. Il s'agit de l'involution linéaire de E qui à tout vecteur $v = w+u$ de E avec $w \in F$ et $u \in F^\perp$ associe le vecteur $w-u$. On note s_F cette application.*

Remarque 2.4.13.
- "Involution linéaire de E" signifie que s_F est une application linéaire de E dans E telle que $s_F \circ s_F = \mathrm{Id}_E$. Autrement dit, s_F est un automorphisme linéaire de E d'inverse lui-même.
- Comme pour toute symétrie vectorielle, on a $s_F = 2p_F - \mathrm{Id}_E$ (si $v = w+u \in E$ avec $w \in F$ et $u \in F^\perp$, on a $(2p_F - \mathrm{Id}_E)(v) = 2w-(w+u) = w-u = s_F(v)$).
- Si $\{e_1,\dots,e_p\}$ est une base de F et $\{e_{p+1},\dots,e_n\}$ est une base de $F^\perp$ alors $\mathcal{B} := \{e_1,\dots,e_p,e_{p+1},\dots,e_n\}$ est une base de E dans laquelle la matrice représentative de s_F est

$$\mathrm{Mat}_{\mathcal{B}}(s_F) = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & -1 & & \\ & & & & \ddots & \\ & & & & & -1 \end{pmatrix} = \begin{pmatrix} I_p & 0_{p,n-p} \\ 0_{n-p,p} & -I_{n-p} \end{pmatrix}.$$

2.5 Représentation matricielle du produit scalaire

Afin d'aider aux calculs, on cherche à représenter le produit scalaire de façon matricielle.

Soit $(E, \langle\cdot,\cdot,\rangle)$ un espace euclidien. Le point de départ pour définir une telle représentation est le suivant. Supposons que E soit de dimension $n \in \mathbb{N}\backslash\{0\}$ et soit $\mathcal{B} = \{e_1, \ldots, e_n\}$ une base (quelconque) de E. On a vu au début de la section précédente que, si v et w sont deux vecteurs de E de coordonnées respectives $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ dans la base $\mathcal{B}$, alors

$$\langle v, w\rangle = \sum_{1\leqslant i,j\leqslant n} x_i y_j \langle e_i, e_j\rangle.$$

Écrivons $X := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \mathrm{Mat}_{\mathcal{B}}(v)$, $Y := \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \mathrm{Mat}_{\mathcal{B}}(w)$ et A la matrice $\big(\langle e_i, e_j\rangle\big)_{1\leqslant i,j\leqslant n}$. On a alors

$$\begin{aligned}
\langle v, w\rangle &= \sum_{1\leqslant i,j\leqslant n} x_i y_j \langle e_i, e_j\rangle \\
&= \sum_{1\leqslant i,j\leqslant n} x_i \langle e_i, e_j\rangle y_j \\
&= \sum_{i=1}^{n} x_i \left(\sum_{j=1}^{n} \langle e_i, e_j\rangle y_j\right) \\
&= \sum_{i=1}^{n} x_i (AY)_i \ \text{(où } (AY)_i \text{ désigne la } i^{\text{ème}} \text{ coordonnée du vecteur colonne } AY) \\
&= {}^t X A Y
\end{aligned}$$

Cette écriture motive la définition suivante :

Définition 2.5.1. *On appelle matrice représentative du produit scalaire $\langle\cdot,\cdot\rangle$ dans la base $\mathcal{B}$ la matrice*

$$\mathrm{Mat}^{\mathrm{ps}}_{\mathcal{B}}\big(\langle\cdot,\cdot\rangle\big) := \big(\langle e_i, e_j\rangle\big)_{1\leqslant i,j\leqslant n} = \begin{pmatrix} \langle e_1, e_1\rangle & \cdots & \langle e_1, e_n\rangle \\ \vdots & & \vdots \\ \langle e_n, e_1\rangle & \cdots & \langle e_n, e_n\rangle \end{pmatrix}$$

Exemple 2.5.2.

- Si $\mathcal{B}_{\mathrm{can}}$ désigne la base canonique de $\mathbb{R}^n$, on a $\mathrm{Mat}^{\mathrm{ps}}_{\mathcal{B}_{\mathrm{can}}}\big(\langle\cdot,\cdot\rangle_{\mathrm{can}}\big) = I_n$.
- On a $\mathrm{Mat}^{\mathrm{ps}}_{\mathcal{B}}\big(\langle\cdot,\cdot\rangle_{\mathcal{B}}\big) = I_n$.

- Considérons sur l'espace vectoriel $\mathbb{R}_2[X]$ de dimension 3 le produit scalaire $\langle\cdot,\cdot\rangle$ défini dans l'exemple 2.2.2 4.. On note $\mathcal{B}$ la base $\{1, X, X^2\}$ de $\mathbb{R}_2[X]$ et on calcule alors

$$\begin{aligned}
\langle 1, 1\rangle &= \int_0^1 1\,\mathrm{d}t = [t]_0^1 = 1\\
\langle 1, X\rangle = \langle X, 1\rangle &= \int_0^1 t\,\mathrm{d}t = \left[\frac{t^2}{2}\right]_0^1 = \frac{1}{2}\\
\left\langle 1, X^2\right\rangle = \left\langle X^2, 1\right\rangle &= \int_0^1 t^2\,\mathrm{d}t = \left[\frac{t^3}{3}\right]_0^1 = \frac{1}{3}\\
\langle X, X\rangle &= \int_0^1 t^2\,\mathrm{d}t = \left[\frac{t^3}{3}\right]_0^1 = \frac{1}{3}\\
\left\langle X, X^2\right\rangle = \left\langle X^2, X\right\rangle &= \int_0^1 t^3\,\mathrm{d}t = \left[\frac{t^4}{4}\right]_0^1 = \frac{1}{4}\\
\left\langle X^2, X^2\right\rangle &= \int_0^1 t^4\,\mathrm{d}t = \left[\frac{t^5}{5}\right]_0^1 = \frac{1}{5}
\end{aligned}$$

Ainsi,

$$\mathrm{Mat}_{\mathcal{B}}^{\mathrm{ps}}\left(\langle\cdot,\cdot\rangle\right) = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3}\\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4}\\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5}\end{pmatrix}$$

Remarque 2.5.3.
- Attention à ne pas faire de confusions avec la matrice représentative d'une application linéaire !

- Avec les notations du début de cette section, on a

$$\langle v, w\rangle = {}^t\mathrm{Mat}_{\mathcal{B}}(v)\mathrm{Mat}_{\mathcal{B}}^{\mathrm{ps}}\big(\langle\cdot,\cdot\rangle\big)\mathrm{Mat}_{\mathcal{B}}(w),$$

qui est aussi égal à ${}^t\mathrm{Mat}_{\mathcal{B}}(w)\mathrm{Mat}_{\mathcal{B}}^{\mathrm{ps}}\big(\langle\cdot,\cdot\rangle\big)\mathrm{Mat}_{\mathcal{B}}(v) = \langle w, v\rangle$ car $\langle\cdot,\cdot\rangle$ est symétrique.

- Comme le produit scalaire est symétrique, on a, pour tous $i, j \in \{1, \ldots, n\}$, $\langle e_i, e_j\rangle = \langle e_j, e_i\rangle$ et la matrice $\mathrm{Mat}_{\mathcal{B}}^{\mathrm{ps}}\big(\langle\cdot,\cdot\rangle\big)$ est donc symétrique i.e. ${}^t\mathrm{Mat}_{\mathcal{B}}^{\mathrm{ps}}\big(\langle\cdot,\cdot\rangle\big) = \mathrm{Mat}_{\mathcal{B}}^{\mathrm{ps}}\big(\langle\cdot,\cdot\rangle\big)$.

  Le fait que le produit scalaire $\langle\cdot,\cdot\rangle$ soit "défini" s'exprime dans le fait que la matrice $\mathrm{Mat}_{\mathcal{B}}^{\mathrm{ps}}\left(\langle\cdot,\cdot\rangle\right)$ est inversible. En effet, si on note $A := \mathrm{Mat}_{\mathcal{B}}^{\mathrm{ps}}\left(\langle\cdot,\cdot\rangle\right)$ et si X est un vecteur colonne de taille n quelconque tel que AX est le vecteur colonne nul de taille n, alors ${}^tXAX = 0$ i.e. $\langle v, v\rangle = 0$ où v désigne le vecteur de coordonnées X dans la base $\mathcal{B}$. Par suite, $v = 0_E$ et X est donc le vecteur colonne nul. Comme le noyau de la matrice carrée A est réduit au vecteur colonne nul, A est inversible.

- La base $\mathcal{B}$ est orthonormale par rapport au produit scalaire $\langle\cdot,\cdot\rangle$ si et seulement si $\mathrm{Mat}_{\mathcal{B}}^{\mathrm{ps}}\big(\langle\cdot,\cdot\rangle\big) = I_n$. Ainsi, le corollaire 2.3.10 affirme qu'il existe toujours une base de E dans laquelle la matrice représentative du produit scalaire $\langle\cdot,\cdot\rangle$ est la matrice identité.

On peut à présent se demander comment sont reliées les matrices représentatives de $\langle\cdot,\cdot\rangle$ dans deux bases (quelconques) différentes, autrement dit s'intéresser à la question du changement de base pour la matrice représentative d'un produit scalaire.

Soit donc $\mathcal{B}'$ une autre base de E et considérons la matrice de passage $P_{\mathcal{B}\to\mathcal{B}'}$ de la base $\mathcal{B}$ à la base $\mathcal{B}'$. On a l'égalité suivante :

Proposition 2.5.4. *On a*

$$\mathrm{Mat}^{\mathrm{ps}}_{\mathcal{B}'}\big(\langle\cdot,\cdot\rangle\big) = {}^tP_{\mathcal{B}\to\mathcal{B}'}\,\mathrm{Mat}^{\mathrm{ps}}_{\mathcal{B}}\big(\langle\cdot,\cdot\rangle\big)P_{\mathcal{B}\to\mathcal{B}'}$$

Démonstration. Soient $v,w\in E$. Comme au début de la section, notons $X := \mathrm{Mat}_{\mathcal{B}}(v)$, $Y := \mathrm{Mat}_{\mathcal{B}}(w)$ et $A := \mathrm{Mat}^{\mathrm{ps}}_{\mathcal{B}}\big(\langle\cdot,\cdot\rangle\big)$. Notons ensuite $X' := \mathrm{Mat}_{\mathcal{B}'}(v)$, $Y' := \mathrm{Mat}_{\mathcal{B}'}(w)$ et $A' := \mathrm{Mat}^{\mathrm{ps}}_{\mathcal{B}'}\big(\langle\cdot,\cdot\rangle\big)$.

On a $X' = P_{\mathcal{B}'\to\mathcal{B}}X = {P_{\mathcal{B}\to\mathcal{B}'}}^{-1}X \Leftrightarrow X = P_{\mathcal{B}\to\mathcal{B}'}X'$ et $Y' = P_{\mathcal{B}'\to\mathcal{B}}Y = {P_{\mathcal{B}\to\mathcal{B}'}}^{-1}Y \Leftrightarrow Y = P_{\mathcal{B}\to\mathcal{B}'}Y'$. Ainsi,

$$\begin{array}{rcl}\langle v,w\rangle & = & {}^tXAY \\ & = & {}^t\big(P_{\mathcal{B}\to\mathcal{B}'}X'\big)\,A\,\big(P_{\mathcal{B}\to\mathcal{B}'}Y'\big) \\ & = & {}^tX'\left({}^tP_{\mathcal{B}\to\mathcal{B}'}A\,P_{\mathcal{B}\to\mathcal{B}'}\right)Y'.\end{array}$$

Remarquons maintenant que, si $M = (m_{ij})_{1\leqslant i,j\leqslant n}$ est une matrice carrée de taille n quelconque et si, pour $i\in\{1,\dots,n\}$, X_i désigne le vecteur colonne avec coordonnées 1 à la ligne i et 0 sur les autres lignes, on a, pour tous $i,j\in\{1,\dots,n\}$, ${}^tX_iMX_j = m_{ij}$.

Soient alors $i,j\in\{1,\dots,n\}$ et notons $\mathcal{B}' = \{e'_1,\dots,e'_n\}$. On a $\mathrm{Mat}_{\mathcal{B}'}(e'_i) = X_i$ et $\mathrm{Mat}_{\mathcal{B}'}\left(e'_j\right) = X_j$ et, par l'égalité ci-dessus,

$$\left\langle e'_i, e'_j\right\rangle = {}^tX_i\left({}^tP_{\mathcal{B}\to\mathcal{B}'}A\,P_{\mathcal{B}\to\mathcal{B}'}\right)X_j.$$

Autrement dit, le coefficient à la ligne i et la colonne j de la matrice A' est égal au coefficient à la ligne i et la colonne j de la matrice ${}^tP_{\mathcal{B}\to\mathcal{B}'}A\,P_{\mathcal{B}\to\mathcal{B}'}$. Par conséquent, $A' = {}^tP_{\mathcal{B}\to\mathcal{B}'}A\,P_{\mathcal{B}\to\mathcal{B}'}$. □

Remarque 2.5.5. Attention à ne surtout pas confondre ce changement de base pour les produits scalaires avec le changement de base pour les applications linéaires.

2.6 Endomorphismes orthogonaux et matrices orthogonales

Soit $(E,\langle\cdot,\cdot\rangle)$ un espace euclidien de dimension $n\in\mathbb{N}\backslash\{0\}$ et soit f un endomorphisme de E.

Définition 2.6.1. *On dit que f est un endomorphisme orthogonal de E si pour tous $v,w\in E$, $\langle f(v),f(w)\rangle = \langle v,w\rangle$.*

Exemple 2.6.2. L'identité Id_E de E est un endomorphisme orthogonal.

Autrement dit, f est un endomorphisme orthogonal de E si f conserve le produit scalaire $\langle\cdot,\cdot\rangle$. Cette propriété est équivalente à la préservation de la norme euclidienne et à la préservation de la distance euclidienne :

Proposition 2.6.3. *f est un endomorphisme orthogonal si et seulement si f conserve la norme euclidienne $\|\cdot\|$ associée à $\langle\cdot,\cdot\rangle$ si et seulement si f conserve la distance euclidienne d associée à $\langle\cdot,\cdot\rangle$.*

Démonstration. On montre que les assertions

1. f est orthogonal,
2. pour tout $v \in E$, $\|f(v)\| = \|v\|$,
3. pour tous $v, w \in E$, $d(f(v), f(w)) = d(v, w)$,

sont équivalentes.

Montrons tout d'abord $1 \Rightarrow 2$: si f préserve le produit scalaire $\langle\cdot,\cdot\rangle$, alors, pour tout $v \in E$,

$$\|f(v)\| = \sqrt{\langle f(v), f(v)\rangle} = \sqrt{\langle v, v\rangle} = \|v\|.$$

Montrons ensuite $2 \Rightarrow 3$: si f préserve la norme euclidienne associé à $\langle\cdot,\cdot\rangle$, alors, pour tous $v, w \in E$,

$$d(f(v), f(w)) = \|f(w) - f(v)\| = \|f(w - v)\| = \|w - v\| = d(v, w).$$

Enfin, on a l'implication $3 \Rightarrow 2$ car, si f préserve la distance euclidienne et si $v \in E$, alors

$$\|f(v)\| = d(f(v), 0) = d(f(v), f(0)) = d(v, 0) = \|v\|,$$

ainsi que l'implication $2 \Rightarrow 1$ car, si f préserve la norme euclidienne et si $v, w \in E$, alors

$$\begin{aligned}
\langle f(v), f(w)\rangle &= \frac{1}{2}\left(\|f(v) + f(w)\|^2 - \|f(v)\|^2 - \|f(w)\|^2\right) \\
&= \frac{1}{2}\left(\|f(v + w)\|^2 - \|f(v)\|^2 - \|f(w)\|^2\right) \\
&= \frac{1}{2}\left(\|v + w\|^2 - \|v\|^2 - \|w\|^2\right) \\
&= \langle v, w\rangle.
\end{aligned}$$

□

Remarque 2.6.4. Un endomorphisme orthogonal est également appelé isométrie.

Une autre caractérisation d'un endomorphisme orthogonal est qu'il transforme une base orthonormale en une base orthonormale :

Proposition 2.6.5. *L'endomorphisme f est orthogonal si et seulement si f associe à toute base orthonormale de E une base orthonormale de E.*

Démonstration. Supposons que f est orthogonal et soit $\{e_1, \dots, e_n\}$ une base orthonormale de E. Comme f est orthogonal, on a, pour tous $i, j \in \{1, \dots, n\}$, $\langle f(e_i), f(e_j)\rangle = \langle e_i, e_j\rangle = \delta_{ij}$, donc la famille de n vecteurs $\{f(e_1), \dots, f(e_n)\}$ est orthonormale : il s'agit donc d'une base orthonormale de E.

Réciproquement, soit $\mathcal{B} = \{e_1, \dots, e_n\}$ une base orthonormale de E et supposons que la famille $\{f(e_1), \dots, f(e_n)\}$ est orthonormale. Soient alors $v, w \in E$, de coordonnées respectives $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ dans la base $\mathcal{B}$, on a

$$\langle f(v), f(w)\rangle = \left\langle f\left(\sum_{i=1}^n x_i e_i\right), f\left(\sum_{j=1}^n y_j e_j\right)\right\rangle = \sum_{1\leqslant i,j\leqslant n} x_i y_j \langle f(e_i), f(e_j)\rangle = \sum_{i=1}^n x_i y_i = \langle v, w\rangle.$$

L'endomorphisme f est donc orthogonal. □

Remarque 2.6.6. Remarquons que, d'après la démonstration ci-dessus, il suffit qu'il existe une base orthonormale $\{e_1, \dots, e_n\}$ de E telle que la famille $\{f(e_1), \dots, f(e_n)\}$ soit orthonormale pour que l'endomorphisme f soit orthogonal.

Corollaire 2.6.7. *Tout endomorphisme orthogonal est bijectif et son inverse est également orthogonal.*

Démonstration. Supposons que l'endomorphisme f est orthogonal et soit $\mathcal{B}$ une base orthonormale de E. Alors, d'après la proposition précédente, la famille $\mathcal{B}' := f(\mathcal{B})$ est une base orthonormale de E. On peut donc considérer la matrice de passage $P_{\mathcal{B}\to\mathcal{B}'}$ qui est également la matrice représentative de f dans la base $\mathcal{B}$ de E : en tant que matrice de passage, cette matrice est inversible et donc l'endomorphisme f est bijectif.

Montrons enfin que l'endomorphisme f^{-1} est également orthogonal : soient $v, w \in E$, alors, comme f est orthogonal,

$$\left\langle f^{-1}(v), f^{-1}(w)\right\rangle = \left\langle f\left(f^{-1}(v)\right), f\left(f^{-1}(w)\right)\right\rangle = \langle v, w\rangle.$$

□

Intéressons-nous maintenant à la caractérisation matricielle de l'"orthogonalité" d'un endomorphisme. Soit $\mathcal{B}$ une base orthonormale de E.

Proposition 2.6.8. *Soit $A := \mathrm{Mat}_{\mathcal{B}}(f)$ la matrice représentative de f dans la base $\mathcal{B}$. L'endomorphisme f de E est orthogonal si et seulement si ${}^tAA = I_n$.*

Démonstration. Supposons que f est orthogonal. Alors, si l'on note $\mathcal{B} = \{e_1, \dots, e_n\}$, on a, pour tous $i, j \in \{1, \dots, n\}$,

$${}^tX_i \left({}^tAA\right) X_j = {}^t(AX_i)(AX_j) = \langle f(e_i), f(e_j)\rangle = \langle e_i, e_j\rangle = \delta_{ij},$$

et donc ${}^tAA = I_n$ (cf. preuve de la proposition 2.5.4).

Réciproquement, supposons que ${}^tAA = I_n$ et soient $v, w \in E$. En notant $X := \mathrm{Mat}_{\mathcal{B}}(v)$ et $Y := \mathrm{Mat}_{\mathcal{B}}(w)$, on a (comme la base $\mathcal{B}$ est orthonormale : remarque 2.5.3)

$$\langle f(v), f(w) \rangle = {}^t(AX)\,(AY) = {}^tX\,{}^tAAY = {}^tXY = \langle v, w \rangle.$$

□

Ce résultat motive la définition suivante :

Définition 2.6.9. *On dit qu'une matrice A de $\mathrm{M}_n(\mathbb{R})$ est orthogonale si ${}^tAA = I_n$.*

Remarque 2.6.10.

- Une matrice orthogonale est inversible et son inverse est sa transposée.
- L'endomorphisme f de E est orthogonal si et seulement si la matrice $\mathrm{Mat}_{\mathcal{B}}(f)$ est orthogonale. "Réciproquement", une matrice A de $\mathrm{M}_n(\mathbb{R})$ est orthogonale si et seulement si l'endomorphisme de $\mathbb{R}^n$ représenté par f dans la base canonique est orthogonal.
- Si f est orthogonal et si A désigne la matrice représentative de f dans la base orthonormale $\mathcal{B}$ de E alors $\mathrm{Mat}_{\mathcal{B}}\left(f^{-1}\right) = {}^tA$:

  $$\mathrm{Mat}_{\mathcal{B}}\left(f^{-1}\right) = \mathrm{Mat}_{\mathcal{B}}\left(f\right)^{-1} = A^{-1} = {}^tA$$

  (la matrice A est orthogonale).

Exemple 2.6.11.

1. La matrice

   $$A := \frac{1}{3}\begin{pmatrix} 2 & -1 & 2 \\ 2 & 2 & -1 \\ -1 & 2 & 2 \end{pmatrix}$$

   est orthogonale.

2. Une symétrie orthogonale est un endomorphisme orthogonal (remarque 2.4.13). Une projection orthogonale sur un sous-espace vectoriel strict n'est pas un endomorphisme orthogonal (remarque 2.4.7).

Un point de vue supplémentaire sur l'égalité "${}^tAA = I_n$" est le suivant : une matrice $A \in \mathrm{M}_n(\mathbb{R})$ est orthogonale si et seulement si les vecteurs colonnes la composant forment une base orthonormale de $\mathbb{R}^n$ muni du produit scalaire canonique. Dans ce cas, la matrice A est la matrice de passage de la base canonique de $\mathbb{R}^n$ à la base orthonormale formée par ses vecteurs colonnes. On peut généraliser cela de la façon suivante (rappelons que la base $\mathcal{B}$ a été supposée orthonormale) :

Proposition 2.6.12. *Soit $\mathcal{B}'$ une base de E. Alors $\mathcal{B}'$ est orthonormale si et seulement si la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$ est orthogonale.*

Démonstration. La matrice de passage $P := P_{\mathcal{B}\to\mathcal{B}'}$ de $\mathcal{B} = \{e_1, \ldots, e_n\}$ à $\mathcal{B}' = \{e'_1, \ldots, e'_n\}$ est formée, dans l'ordre, des vecteurs coordonnées des vecteurs e'_i, $i = 1, \ldots, n$, dans la base $\mathcal{B}$. On a alors, comme $\mathcal{B}$ est orthonormale,

$$^tPP = \begin{pmatrix} \langle e'_1, e'_1\rangle & \cdots & \langle e'_1, e'_n\rangle \\ \vdots & & \vdots \\ \langle e'_n, e'_1\rangle & \cdots & \langle e'_n, e'_n\rangle \end{pmatrix}$$

Ainsi, la matrice de passage $P = P_{\mathcal{B}\to\mathcal{B}'}$ est orthogonale ssi la famille de vecteurs $\{e'_1, \ldots, e'_n\}$ est orthonormale. □

Remarque 2.6.13. En particulier, toute matrice de passage d'une base orthonormale à une autre est orthogonale.

Remarquons ensuite que le déterminant d'une matrice orthogonale, et donc d'un endormorphisme orthogonal, est égal à 1 ou -1 :

Proposition et Définition 2.6.14. *Soit A une matrice orthogonale de* $\mathrm{M}_n(\mathbb{R})$. *Alors* $\det(A) = 1$ *ou* $\det(A) = -1$. *Ainsi, si f est orthogonal,* $\det(f) = 1$ *ou* $\det(f) = -1$.

Si $\det(A) = 1$*, resp.* $\det(f) = 1$*, on dit que A, resp. f, est une matrice orthogonale <u>directe</u>, resp. endomorphisme orthogonal <u>direct</u>. Si* $\det(A) = -1$*, resp.* $\det(f) = -1$*, on dit que A, resp. f, est une matrice orthogonale <u>indirecte</u>, resp. endomorphisme orthogonal <u>indirect</u>.*

Démonstration. On a $^tAA = I_n$ donc $1 = \det\left({}^tAA\right) = \det\left({}^tA\right)\det(A) = (\det(A))^2$ d'où $\det(A) = 1$ ou $\det(A) = -1$

Ainsi, si f est orthogonal, comme sa matrice représentative dans une base orthonormale est orthogonale (proposition 2.6.8), on a $\det(f) = 1$ ou $\det(f) = -1$. □

Exemple 2.6.15.
1. La matrice orthogonale A de l'exemple 2.6.11 1. est directe.
2. Une symétrie orthogonale par rapport à un sous-espace vectoriel F de E est directe ou indirecte suivant la parité de la "codimension" de F : le déterminant d'une telle symétrie orthogonale est $(-1)^{n-p}$ si p est la dimension de F (remarque 2.4.13).

Terminons cette section en remarquant que l'ensemble des endomorphismes orthogonaux de E muni de la composition forme un groupe. On a déjà vu plus haut que l'identité Id_E de E est orthogonale et que, si f est un endomorphisme orthogonal, f est bijectif et son inverse est également orthogonal (corollaire 2.6.7). Enfin, la composition d'endomorphismes orthogonaux est également orthogonale :

Proposition 2.6.16. *Soit g un endomorphisme de E et supposons que f et g sont orthogonaux. Alors la composition $f \circ g$ (ainsi que la composition $g \circ f$) est orthogonale.*

Démonstration. Soient $x, y \in E$, on a

$$\begin{aligned} \left\langle (g\circ f)(v), (g\circ f)(w)\right\rangle &= \left\langle g\big(f(v)\big), g\big(f(w)\big)\right\rangle \\ &= \langle f(v), f(w)\rangle \text{ (car } g \text{ est orthogonal)} \\ &= \langle v, w\rangle \text{ (car } f \text{ est orthogonal).} \end{aligned}$$

□

Corollaire et Définition 2.6.17. *L'ensemble, noté $\mathcal{O}(E, \langle\cdot,\cdot\rangle)$ ou simplement $\mathcal{O}(E)$ lorsque le contexte est clair, des endomorphismes orthogonaux de l'espace euclidien $(E, \langle\cdot,\cdot\rangle)$ est un sous-groupe du groupe $(\mathcal{GL}(E), \circ)$ des automorphismes linéaires de E muni de la composition (appelé groupe linéaire de E). On appelle $(\mathcal{O}(E), \circ)$ le groupe orthogonal de $(E, \langle\cdot,\cdot\rangle)$.*

De manière équivalente, l'ensemble $\mathrm{O}_n(\mathbb{R})$ des matrices orthogonales de $\mathrm{M}_n(\mathbb{R})$ est un sous-groupe du groupe $(\mathrm{GL}_n(\mathbb{R}), \cdot)$ des matrices réelles inversibles de taille n muni du produit matriciel. $\mathrm{O}_n(\mathbb{R})$ est appelé groupe orthogonal de $\mathrm{M}_n(\mathbb{R})$.

Remarque 2.6.18. Le sous-ensemble de $\mathcal{O}(E)$ constitué des endomorphismes orthogonaux directs de E est un sous-groupe de $(\mathcal{O}(E), \circ)$, appelé groupe spécial orthogonal de E et noté $\mathcal{SO}(E)$.

Le sous-ensemble de $\mathrm{O}_n(\mathbb{R})$ constitué des matrices orthogonales directes de taille n est un sous-groupe de $(\mathrm{O}_n(\mathbb{R}), \cdot)$, appelé groupe spécial orthogonal de $\mathrm{M}_n(\mathbb{R})$ et noté $\mathrm{SO}_n(\mathbb{R})$.

2.7 Décomposition QR d'une matrice inversible

Dans cette section, nous allons étudier le pendant matriciel du procédé d'orthonormalisation de Gram-Schmidt (théorème 2.3.7). Précisément, soit $n \in \mathbb{N}\backslash\{0\}$ et considérons une matrice A de $\mathrm{GL}_n(\mathbb{R})$. Notons $v_1, \ldots, v_n$ les vecteurs colonnes qui, dans l'ordre, forment la matrice A. La famille $\mathcal{B} := \{v_1, \ldots, v_n\}$, considérée comme famille de vecteurs de $\mathbb{R}^n$, est une base de $\mathbb{R}^n$: si $\mathcal{B}_0$ désigne la base canonique de $\mathbb{R}^n$, A est alors la matrice de passage $P_{\mathcal{B}_0 \to \mathcal{B}}$. Notons ensuite $\mathcal{B}' = \{e_1, \ldots, e_n\}$ la base orthonormale de $\mathbb{R}^n$ (par rapport au produit scalaire canonique) obtenue à l'aide du procédé d'orthonormalisation de Gram-Schmidt appliqué à la base $\mathcal{B} = \{v_1, \ldots, v_n\}$. On note $P := P_{\mathcal{B} \to \mathcal{B}'}$ la matrice de passage de la base $\{v_1, \ldots, v_n\}$ à la base $\{e_1, \ldots, e_n\}$ de $\mathbb{R}^n$.

Dans le procédé de Gram-Schmidt, pour $k \in \{1, \ldots, n\}$, le vecteur e_k est défini comme une combinaison linéaire des vecteurs $\{v_1, \ldots, v_k\}$. En particulier, la matrice de passage P est triangulaire supérieure. De par cette construction, on sait également que les coefficients diagonaux de P sont strictement positifs.

Enfin, si Q désigne la matrice de passage de $\mathcal{B}_0$ à $\mathcal{B}'$ (i.e. la matrice formée, dans l'ordre, par les vecteurs colonnes $e_1, \ldots, e_n$), on a $Q = AP$ ($P_{\mathcal{B}_0 \to \mathcal{B}'} = P_{\mathcal{B}_0 \to \mathcal{B}} P_{\mathcal{B} \to \mathcal{B}'}$). On réécrit cette égalité sous la forme $A = QR$ avec $R := P^{-1} = P_{\mathcal{B}' \to \mathcal{B}}$. Remarquons enfin que, comme les vecteurs colonnes de la matrice Q forment une base orthonormale de $\mathbb{R}^n$ (par rapport au produit scalaire canonique), Q est orthogonale.

Au total, nous avons (partiellement) montré le résultat suivant :

Théorème 2.7.1 (Décomposition QR d'une matrice inversible)**.** *Soit $n \in \mathbb{N}\backslash\{0\}$ et soit $A \in \mathrm{GL}_n(\mathbb{R})$. Il existe une unique matrice orthogonale $Q \in \mathrm{O}_n(\mathbb{R})$ et une unique matrice $R \in \mathrm{GL}_n(\mathbb{R})$ triangulaire supérieure à coefficients diagonaux strictement positifs telles que $A = QR$. On appelle cette écriture la décomposition QR de A.*

Démonstration. Il nous reste à montrer l'unicité de la décomposition. Soient donc $Q_1, Q_2 \in \mathrm{O}_n(\mathbb{R})$ et R_1, R_2 deux matrices triangulaires supérieures de $\mathrm{GL}_n(\mathbb{R})$ à coefficients diagonaux strictement positifs telles que $A = Q_1 R_1 = Q_2 R_2$.

On a alors $R_1 R_2^{-1} = Q_1^{-1} Q_2$ et $R_1 R_2^{-1} = Q_1^{-1} Q_2$ est alors une matrice orthogonale ($(\mathrm{O}_n(\mathbb{R}), \cdot)$ est un groupe) et triangulaire supérieure à coefficients strictement positifs (l'inverse d'une matrice triangulaire supérieure à coefficients strictement positifs est une matrice triangulaire supérieure à coefficients strictement positifs). Il s'agit donc de la matrice I_n d'après le lemme qui suit cette preuve.

Ainsi, $R_1 R_2^{-1} = Q_1^{-1} Q_2 = I_n$ et donc $Q_1 = Q_2$ et $R_1 = R_2$. □

Lemme 2.7.2. *Soit $A \in \mathrm{GL}_n(\mathbb{R})$ une matrice orthogonale et triangulaire supérieure à coefficients strictement positifs. Alors $A = I_n$.*

Démonstration. Notons $v_1, \ldots, v_n$ les vecteurs colonnes qui, dans l'ordre, forment la matrice $A = (a_{i\,j})_{1 \leqslant i,j \leqslant n}$. Le fait que A soit orthogonale signifie que, pour tout $i, j \in \{1, \ldots, n\}$, $\langle v_i, v_j \rangle = \delta_{i\,j}$. En particulier, $\langle v_1, v_1 \rangle = 1$. Mais, comme A est triangulaire supérieure, $\langle v_1, v_1 \rangle = a_{1\,1}^2$ et, comme $a_{1\,1}$ est strictement positif, on a nécessairement $a_{1\,1} = 1$. Pour tout $j \in \{2, \ldots, n\}$, on a alors $\langle v_1, v_j \rangle = a_{1\,j} = 0$ (autrement dit, la première ligne n'a que des coefficients nuls sauf le coefficient $a_{1\,1}$ qui est égal à 1).

Soit $k \in \{1, n-1\}$ et supposons que l'on a déjà montré que, pour tout $i \in \{1, \ldots, k\}$ et tout $j \in \{1, \ldots, n\}$, $a_{i\,j} = \delta_{i\,j}$ (autrement dit que pour les k premières lignes, tous les coefficients d'une ligne i donnée sont nuls sauf le coefficient $a_{i\,i}$ qui est égal à 1). On considère alors le vecteur colonne v_{k+1} dont, par hypothèse de récurrence, la seule coordonnée non nulle est $a_{k+1\,k+1} > 0$. Comme $\langle v_{k+1}, v_{k+1} \rangle = {a_{k+1\,k+1}}^2 = 1$, on a $a_{k+1\,k+1} = 1$ et, par suite, pour tout $j \in \{k+2, \ldots, n\}$, $\langle v_{k+1}, v_j \rangle = a_{k+1\,j} = 0$: la ligne $k+1$ n'a donc que des coefficients nuls sauf le coefficient $a_{k+1\,k+1}$ qui est égal à 1. □

Exemple 2.7.3. Considérons la matrice

$$A := \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

de $\mathrm{M}_3(\mathbb{R})$. On note $v_1 := (1,1,0)$, $v_2 := (1,2,0)$, $v_3 := (0,0,1) \in \mathbb{R}^3$ et on applique le procédé d'orthonormalisation de Gram-Schmidt à la base $\{v_1, v_2, v_3\}$ de $\mathbb{R}^3$. On pose

$$\begin{aligned}
\epsilon_1 &:= v_1 = (1,1,0) \\
\epsilon_2 &:= v_2 - \frac{\langle v_2, \epsilon_1 \rangle}{\|\epsilon_1\|^2}\epsilon_1 = v_2 - \frac{3}{2}\epsilon_1 = \left(-\frac{1}{2}, \frac{1}{2}, 0\right) = \frac{1}{2}(-1,1,0) \\
\epsilon_2' &:= (-1,1,0) \\
\epsilon_3 &:= v_3 - \frac{\langle v_3, \epsilon_1 \rangle}{\|\epsilon_1\|^2}\epsilon_1 - \frac{\langle v_3, \epsilon_2' \rangle}{\|\epsilon_2'\|^2}\epsilon_2' = v_3 = (0,0,1) \\
e_1 &:= \frac{1}{\|\epsilon_1\|}\epsilon_1 = \frac{1}{\sqrt{2}}(1,1,0) \\
e_2 &:= \frac{1}{\|\epsilon_2'\|}\epsilon_2' = \frac{1}{\sqrt{2}}(-1,1,0) \\
e_3 &:= \frac{1}{\|\epsilon_3\|}\epsilon_3 = (0,0,1)
\end{aligned}$$

Comme

$$\begin{aligned} v_1 &= \epsilon_1 = \sqrt{2}e_1 \\ v_2 &= \frac{3}{2}\epsilon_1 + \epsilon_2 = \frac{3}{2}\epsilon_1 + \frac{1}{2}\epsilon_2' = \frac{3}{\sqrt{2}}e_1 + \frac{1}{\sqrt{2}}e_2 \\ v_3 &= \epsilon_3 = e_3 \end{aligned}$$

la matrice de passage de la base $\{e_1, e_2, e_3\}$ à la base $\{v_1, v_2, v_3\}$ est

$$R := \begin{pmatrix} \sqrt{2} & \frac{3}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Si on note Q la matrice

$$\begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

formée par les coordonnées (dans la base canonique de $\mathbb{R}^3$) des vecteurs e_1, e_2, e_3, l'égalité

$$A = QR = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{2} & \frac{3}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

est la décomposition QR de A.

Remarque 2.7.4. Avec les notations précédentes, on peut également calculer R en utilisant l'égalité $R = Q^{-1}A = {}^tQA$ (Q est orthogonale).

2.8 Endomorphismes symétriques et matrices symétriques

Soit $(E, \langle\cdot,\cdot\rangle)$ un espace euclidien de dimension $n \in \mathbb{N}\backslash\{0\}$ et soit f un endomorphisme de E.

Définition 2.8.1. *On dit que l'endomorphisme f de E est symétrique si pour tous $v, w \in E$, $\langle f(v), w\rangle = \langle v, f(w)\rangle$.*

L'endomorphisme f est symétrique si et seulement sa matrice dans une base orthonormale de $(E, \langle\cdot,\cdot\rangle)$ est symétrique :

Proposition 2.8.2. *Soit $\mathcal{B}$ une base orthonormale de E et notons $A := \mathrm{Mat}_{\mathcal{B}}(f)$ la matrice représentative de f dans $\mathcal{B}$. Alors f est symétrique si et seulement si la matrice A est symétrique i.e. ${}^tA = A$.*

Démonstration. Supposons que f est symétrique. Alors, si l'on note $\mathcal{B} = \{e_1, \ldots, e_n\}$, on a, pour tous $i, j \in \{1, \ldots, n\}$,

$${}^tX_i{}^tAX_j = {}^t(AX_i)X_j = \langle f(e_i), e_j\rangle = \langle e_i, f(e_j)\rangle = {}^tX_i(AX_j) = {}^tX_iAX_j,$$

et donc ${}^tA = A$.

Réciproquement, supposons que ${}^tA = A$ et soient $v, w \in E$. En notant $X := \mathrm{Mat}_{\mathcal{B}}(v)$ et $Y := \mathrm{Mat}_{\mathcal{B}}(w)$, on a

$$\langle f(v), w\rangle = {}^t(AX)\,Y = {}^tX\,{}^tAY = {}^tXAY = {}^tX(AY) = \langle v, f(w)\rangle\,.$$

□

Exemple 2.8.3. L'endomorphisme $h : \mathbb{R}^2 \to \mathbb{R}^2\ ;\ (x,y) \mapsto (x+2y, 2x+y)$ de $\mathbb{R}^2$ est symétrique par rapport au produit scalaire canonique de $\mathbb{R}^2$.

Une propriété remarquable des endomorphismes symétriques est qu'ils sont diagonalisables, et ce, dans une base orthonormale. Avant d'énoncer ce résultat, donnons la définition suivante : deux sous-espaces vectoriels F_1 et F_2 de E sont dits orthogonaux si pour tous vecteurs v_1 de F_1 et v_2 de F_2, $\langle v_1, v_2\rangle = 0$.

Théorème 2.8.4 (Théorème spectral)**.** *On suppose que l'endomorphisme f est symétrique. Alors*

- *f est diagonalisable,*
- *les sous-espaces propres de f sont orthogonaux deux à deux.*

En particulier, si l'on considère, pour chacun de ces espaces propres, une base orthonormale, la réunion de ces bases est une base orthonormale de E : f est donc diagonalisable dans une base orthonormale de E.

Démonstration. On commence par montrer que le polynôme caractéristique de f est scindé sur $\mathbb{R}$. Pour cela, soit $\mathcal{B}_0$ une base orthonormale de E et considérons la matrice $A := \mathrm{Mat}_{\mathcal{B}_0}(f)$ de f dans $\mathcal{B}_0$. A est une matrice de $\mathrm{M}_n(\mathbb{R})$ et peut être également considérée comme une matrice de $\mathrm{M}_n(\mathbb{C})$. Nous allons montrer que les racines complexes du polynôme caractéristique de A (autrement dit les valeurs propres complexes de A) sont toutes réelles : $\chi_f = \chi_A$ sera donc également scindé sur $\mathbb{R}$.

Soit donc une valeur propre complexe $\lambda \in \mathbb{C}$ de A et montrons que $\lambda \in \mathbb{R}$. Soit $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathrm{M}_{n,1}(\mathbb{C})$ un vecteur propre de A associé à λ : on a donc $AX = \lambda X$. Nous allons appliquer la conjugaison complexe à cette dernière égalité : si $M = (m_{i\,j})_{1\leqslant i\leqslant p, 1\leqslant j\leqslant q}$ est une matrice de $\mathrm{M}_{p,q}(\mathbb{C})$, la matrice conjuguée $\overline{M}$ de M est la matrice $(\overline{m_{i\,j}})_{1\leqslant i,j\leqslant n}$ de $\mathrm{M}_{p,q}(\mathbb{C})$. On obtient alors

$$\overline{AX} = \overline{\lambda X} \Leftrightarrow \overline{A}\,\overline{X} = \overline{\lambda}\,\overline{X} \Leftrightarrow A\overline{X} = \overline{\lambda}\,\overline{X}$$

(les coefficients de A sont réels).

On considère d'autre part l'égalité ${}^t(AX)\overline{X} = {}^tXA\overline{X}$, satisfaite car ${}^tA = A$ (car f est symétrique : proposition 2.8.2). On y remplace AX par λX et $A\overline{X}$ par $\overline{\lambda}\,\overline{X}$ pour obtenir l'égalité ${}^t(\lambda X)\overline{X} = {}^tX\overline{\lambda}\,\overline{X}$ i.e. $\lambda \sum_{i=1}^{n} |x_i|^2 = \overline{\lambda} \sum_{i=1}^{n} |x_i|^2$ et donc $\lambda = \overline{\lambda}$ car X n'est pas le vecteur colonne nul (X est un vecteur propre). Ainsi, $\lambda \in \mathbb{R}$.

Le polynôme caractéristique de l'endomorphisme symétrique f est donc scindé sur $\mathbb{R}$. En particulier, le spectre de f est non-vide.

On montre à présent que f est diagonalisable. On le montre par récurrence sur la dimension n de E. Précisément, on montre par récurrence l'assertion suivante : pour tout $n \in \mathbb{N}\backslash\{0\}$, pour tout espace euclidien $(E, \langle\cdot,\cdot\rangle)$ de dimension n, pour tout endomorphisme symétrique f de E, f est diagonalisable.

Toute matrice carrée de taille 1 étant diagonale, le résultat est vrai pour $n = 1$.

Supposons à présent la propriété vraie au rang $n-1$ pour $n \in \mathbb{N}\backslash\{0; 1\}$ fixé, et montrons-la pour l'endomorphisme symétrique f de l'espace euclidien E de dimension n. Soit $\lambda \in \mathbb{R}$ une valeur propre de f (un tel λ existe car $\mathrm{Sp}_{\mathbb{R}}(f) \neq \varnothing$). Soit ensuite v un vecteur propre de f associé à λ et notons $F := \{v\}^{\perp} = (\mathrm{Vect}\,\{v\})^{\perp}$. Montrons que F est stable par f : soit $w \in F$ alors

$$\begin{array}{rcl} \langle f(w), v\rangle & = & \langle w, f(v)\rangle \ \ (f \text{ est symétrique}) \\ & = & \langle w, \lambda v\rangle \ \ (v \in E_\lambda) \\ & = & \lambda\langle w, v\rangle \\ & = & 0 \ (\text{car } w \in F = \{v\}^{\perp}). \end{array}$$

Ainsi F est stable par f et on peut donc restreindre f en l'endomorphisme $f_{|F} : \begin{array}{rcl} F & \to & F \\ u & \mapsto & f(u) \end{array}$ de F, qui est également symétrique (par rapport à la restriction du produit scalaire $\langle\cdot,\cdot\rangle$ sur F). Comme $\dim(F) = \dim\left((\mathrm{Vect}\,\{v\})^{\perp}\right) = \dim(E) - \dim(\mathrm{Vect}\{v\}) = n-1$, on peut ensuite appliquer l'hypothèse de récurrence à l'endomorphisme symétrique $f_{|F}$ de F : $f_{|F}$ est diagonalisable i.e. il existe une base de F formée de vecteurs propres $e_2, \ldots, e_n$ pour $f_{|F}$. Les sous-espaces vectoriels $\mathrm{Vect}\{v\}$ et F étant en somme directe (proposition 2.4.5), la famille $\{v, e_2, \ldots, e_n\}$ est alors une base de E, formée de vecteurs propres pour f donc f est diagonalisable.

On montre enfin que les sous-espaces propres de f sont deux à deux orthogonaux. Soient donc λ_1 et λ_2 deux valeurs propres distinctes de f et soient $v_1 \in E_{\lambda_1}, v_2 \in E_{\lambda_2}$. Montrons que les vecteurs v_1 et v_2 sont orthogonaux. On a d'une part

$$\langle f(v_1), v_2\rangle = \langle \lambda_1 v_1, v_2\rangle = \lambda_1\langle v_1, v_2\rangle$$

et, d'autre part, comme f est symétrique,

$$\langle f(v_1), v_2\rangle = \langle v_1, f(v_2)\rangle = \langle v_1, \lambda_2 v_2\rangle = \lambda_2\langle v_1, v_2\rangle.$$

Ainsi, $\lambda_1\langle v_1, v_2\rangle = \lambda_2\langle v_1, v_2\rangle$ i.e. $(\lambda_1 - \lambda_2)\langle v_1, v_2\rangle = 0$ et donc $\langle v_1, v_2\rangle = 0$ car $\lambda_1 \neq \lambda_2$. Les vecteurs v_1 et v_2 sont donc orthogonaux. □

Exemple 2.8.5. On considère l'endomorphisme

$$f : \begin{array}{ccc} \mathbb{R}^3 & \to & \mathbb{R}^3 \\ (x, y, z) & \mapsto & (5x - y + 2z, -x + 5y + 2z, 2x + 2y + 2z) \end{array}$$

de $\mathbb{R}^3$. La matrice représentative de f dans la base canonique de $\mathbb{R}^3$ (qui est une base orthonormale pour le produit scalaire canonique de $\mathbb{R}^3$) est

$$A := \begin{pmatrix} 5 & -1 & 2 \\ -1 & 5 & 2 \\ 2 & 2 & 2 \end{pmatrix}.$$

La matrice A étant symétrique, l'endomorphisme f est symétrique.

Le polynôme caractéristique de f est $\chi_f = (6 - X)^2(-X)$ et les valeurs propres de f sont donc 6 et 0.

On a

$$E_6 = \text{Ker } (f - 6\,\text{Id}_{\mathbb{R}^3}) = \left\{(x, y, z) \in \mathbb{R}^3 \mid -x - y + 2z = 0\right\}$$

et

$$E_0 = \text{Ker } f.$$

La famille $\{(2, 0, 1), (1, 1, 0)\}$ est une base de E_6. En appliquant le procédé d'orthonormalisation de Gram-Schmidt à cette famille libre de $\mathbb{R}^3$, on obtient la base orthonormale $\left\{\frac{1}{\sqrt{5}}(2, 0, 1), \frac{1}{\sqrt{30}}(1, -5, -2)\right\}$ de E_6.

D'autre part, le vecteur $\frac{1}{\sqrt{6}}(-1, -1, 2)$ de norme 1 engendre E_0.

Si l'on note $\mathcal{B} := \left\{\frac{1}{\sqrt{5}}(2, 0, 1), \frac{1}{\sqrt{30}}(1, -5, -2), \frac{1}{\sqrt{6}}(-1, -1, 2)\right\}$, la famille $\mathcal{B}$ est alors une base orthonormale de $\mathbb{R}^3$ et la matrice représentative de f dans $\mathcal{B}$ est

$$\begin{pmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Remarque 2.8.6. La réciproque du théorème 2.8.4 est également vraie : si l'endomorphisme f est diagonalisable dans une base orthonormale de E, alors f est symétrique. En effet, soient $\mathcal{B}$ une base orthonormale de E et D une matrice diagonale de $\mathrm{M}_n(\mathbb{R})$ telles que $\mathrm{Mat}_{\mathcal{B}}(f) = D$, alors

$${}^t\mathrm{Mat}_{\mathcal{B}}(f) = {}^tD = D = \mathrm{Mat}_{\mathcal{B}}(f).$$

Le pendant matriciel du théorème 2.8.4 consiste en l'énoncé suivant :

Corollaire 2.8.7. *Soit A une matrice de $\mathrm{M}_n(\mathbb{R})$. On suppose que A est symétrique. Alors il existe une matrice orthogonale $O \in \mathrm{O}_n(\mathbb{R})$ et une matrice diagonale $D \in \mathrm{M}_n(\mathbb{R})$ telles que*

$$D = O^{-1}AO = {}^tOAO$$

Démonstration. A est la matrice représentative d'un endomorphisme h de $\mathbb{R}^n$ dans la base canonique $\mathcal{B}_0$ de $\mathbb{R}^n$, qui est une base orthonormale pour le produit scalaire canonique de $\mathbb{R}^n$. Comme A est symétrique, h est symétrique et, par le théorème précédent, il existe une base orthonormale $\mathcal{B}$ et une matrice diagonale $D \in \mathrm{M}_n(\mathbb{R})$ telle que $\mathrm{Mat}_{\mathcal{B}}(h) = D$.

Les bases $\mathcal{B}_0$ et $\mathcal{B}$ étant des bases orthonormales de $\mathbb{R}^n$, la matrice de passage $O := P_{\mathcal{B}_0 \to \mathcal{B}}$ est une matrice orthogonale (proposition 2.6.12) et on a alors

$$ {}^tOAO = O^{-1}AO = P_{\mathcal{B} \to \mathcal{B}_0} \,\mathrm{Mat}_{\mathcal{B}_0}(h)\, P_{\mathcal{B}_0 \to \mathcal{B}} = \mathrm{Mat}_{\mathcal{B}}(h) = D. $$

□

Exemple 2.8.8. Si l'on reprend la matrice $A := \begin{pmatrix} 5 & -1 & 2 \\ -1 & 5 & 2 \\ 2 & 2 & 2 \end{pmatrix}$ de l'exemple 2.8.5 précédent et si l'on note

$$ O := \begin{pmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{30}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{-5}{\sqrt{30}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{30}} & \frac{-2}{\sqrt{6}} \end{pmatrix}, $$

la matrice O est orthogonale et on a

$$ O^{-1}AO = {}^tOAO = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 0 \end{pmatrix}. $$

2.9 Réduction des endomorphismes et matrices orthogonaux

Soit $(E, \langle\cdot,\cdot\rangle)$ un espace euclidien de dimension $n \in \mathbb{N}\backslash\{0\}$ et soit maintenant f un endomorphisme orthogonal de E. Nous allons montrer le théorème de réduction suivant :

Théorème 2.9.1. *Il existe une base orthonormale $\mathcal{B}$ de E dans laquelle la matrice représentative de f est de la forme*

$$ \mathrm{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} \epsilon_1 & & & & & 0 \\ & \ddots & & & & \\ & & \epsilon_r & & & \\ & & & R(\theta_1) & & \\ & & & & \ddots & \\ 0 & & & & & R(\theta_s) \end{pmatrix} $$

où $r, s \in \mathbb{N}$, pour tout $i \in \{1, \ldots, r\}$, $\epsilon_i \in \{+1; -1\}$ et, pour tout $j \in \{1, \ldots, s\}$, $R(\theta_j) = \begin{pmatrix} \cos\theta_j & -\sin\theta_j \\ \sin\theta_j & \cos\theta_j \end{pmatrix}$ avec $\theta_j \in \mathbb{R}\backslash\{k\pi \mid k \in \mathbb{Z}\}$.

La version matricielle de ce résultat est le suivant : si $A \in \mathrm{O}_n(\mathbb{R})$, il existe une matrice orthogonale $P \in \mathrm{O}_n(\mathbb{R})$ telle que $P^{-1}AP = {}^tPAP$ soit de la forme ci-dessus.

Remarque 2.9.2. En particulier, si l'on applique le théorème 2.9.1 en dimensions 2 et 3, on obtient que :

- Les isométries directes (i.e. les endomorphismes orthogonaux de déterminant 1) de $\mathbb{R}^2$ (muni du produit scalaire canonique) sont les rotations de $\mathbb{R}^2$, chacune de matrice représentative $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$, $\theta \in \mathbb{R}$, dans la base canonique de $\mathbb{R}^2$ (cf. lemme 2.9.4 ci-dessous).

  Les isométries indirectes (i.e. les endomorphismes orthogonaux de déterminant -1) de $\mathbb{R}^2$ (muni du produit scalaire canonique) sont les symétries orthogonales par rapport à une droite vectorielle, chacune de matrice représentative $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ dans une base formée d'un vecteur engendrant la droite vectorielle et d'un vecteur orthogonal au premier (par rapport au produit scalaire canonique de $\mathbb{R}^2$).

- Les isométries directes de $\mathbb{R}^3$ (muni du produit scalaire canonique) sont les rotations autour d'un axe, chacune de matrice $\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$, $\theta \in \mathbb{R}$, dans une base orthonormale adaptée (quand $\theta = \pi$, on parle de <u>retournement</u>).

  Les isométries indirectes de $\mathbb{R}^3$ sont les compositions d'une symétrie orthogonale par rapport à un plan (une telle transformation est également appelée <u>réflexion</u>) et d'une rotation autour de l'axe orthogonal à ce plan (par rapport au produit scalaire canonique de $\mathbb{R}^3$), chacune de matrice représentative $\begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$, $\theta \in \mathbb{R}$, dans une base orthonormale adaptée ($\begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$).

Pour montrer le théorème 2.9.1, nous utiliserons les lemmes suivants :

Lemme 2.9.3. *Soit F un sev de E. Si F est stable par l'endomorphisme orthogonal f, alors l'orthogonal $F^\perp$ de F est également stable par f.*

Démonstration. Soit $v \in F^\perp$. On montre que $f(v) \in F^\perp$. Soit donc $w \in F$. La restriction $f_{|F} : \begin{array}{ccc} F & \to & F \\ u & \mapsto & f(u) \end{array}$ (F est stable par f) est un endomorphisme orthogonal de F (par rapport à la restriction du produit scalaire $\langle \cdot, \cdot \rangle$ sur F), en particulier bijectif : il existe donc $\widetilde{w} \in F$ tel que $w = f(\widetilde{w})$.

On a alors

$$\begin{aligned} \langle f(v), w \rangle &= \langle f(v), f(\widetilde{w}) \rangle \\ &= \langle v, \widetilde{w} \rangle \text{ (f est orthogonal)} \\ &= 0 \text{ (car $v \in F^\perp$ et $\widetilde{w} \in F$).} \end{aligned}$$

□

Lemme 2.9.4. *Soit $A \in \mathrm{M}_2(\mathbb{R})$. Alors A est orthogonale si et seulement si A est de la forme*

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \quad ou \quad \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$$

avec $\theta \in \mathbb{R}$.

Démonstration. Soit $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \in \mathrm{M}_2(\mathbb{R})$. Commençons par remarquer que si A est de l'une des deux formes ci-dessus, alors A est orthogonale. Supposons maintenant que A est orthogonale, i.e. $a^2 + b^2 = 1$, $c^2 + d^2 = 1$ et $ac + bd = 0$. En particulier, les points (a, b) et (c, d) de $\mathbb{R}^2$ appartiennent au cercle de centre 0 et de rayon 1 : il existe donc $\theta, \theta' \in \mathbb{R}$ tels que $a = \cos(\theta)$, $b = \sin(\theta)$, $c = \cos(\theta')$, $d = \sin(\theta')$. Alors

$$\begin{aligned} ac + bd = 0 \quad &\Leftrightarrow \quad \cos(\theta)\cos(\theta') + \sin(\theta)\sin(\theta') = 0 \\ &\Leftrightarrow \quad \cos(\theta - \theta') = 0 \\ &\Leftrightarrow \quad \theta' - \theta = \frac{\pi}{2} + k\pi,\ k \in \mathbb{Z} \\ &\Leftrightarrow \quad \theta' = \theta + \frac{\pi}{2} + k\pi,\ k \in \mathbb{Z}. \end{aligned}$$

Si k est pair, $c = \cos(\theta + \frac{\pi}{2} + k\pi) = \cos(\theta + \frac{\pi}{2}) = -\sin(\theta)$ et $d = \sin(\theta + \frac{\pi}{2} + k\pi) = \sin(\theta + \frac{\pi}{2}) = \cos(\theta)$, donc

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Si k est impair, $c = \cos(\theta + \frac{\pi}{2} + k\pi) = \cos(\theta + \frac{3\pi}{2}) = \sin(\theta)$ et $d = \sin(\theta + \frac{\pi}{2} + k\pi) = \sin(\theta + \frac{3\pi}{2}) = -\cos(\theta)$, donc

$$A = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}.$$

□

Démonstration du théorème 2.9.1. On procède par récurrence sur la dimension n de E : on montre par récurrence que pour tout $n \in \mathbb{N}\backslash\{0\}$, tout espace euclidien $(E, \langle\cdot,\cdot\rangle)$ de dimension n et tout endomorphisme orthogonal f de E, il existe une base orthonormale de E dans laquelle la matrice représentative de f est de la forme voulue.

Pour $n = 1$, soit $E = \mathrm{Vect}\{v\}$ un $\mathbb{R}$-espace vectoriel de dimension 1 muni d'un produit scalaire $\langle\cdot,\cdot\rangle$ et soit f un endomorphisme orthogonal de E. Il existe $\alpha \in \mathbb{R}$ tel que $f(v) = \alpha v$ et alors $\|f(v)\| = \|\alpha v\| = |\alpha|\|v\|$. De plus, comme f est orthogonal, on a $\|f(v)\| = \|v\|$ donc $|\alpha|\|v\| = \|v\|$ et donc $|\alpha| = 1$ ($\|v\| \neq 0$ car $v \neq 0_E$ car v engendre E qui est de dimension 1). Ainsi $\alpha \in \{+1; -1\}$ et la matrice de f dans la base $\{v\}$, ainsi que dans la base orthonormale $\left\{\frac{v}{\|v\|}\right\}$, est (α).

Supposons maintenant la propriété vérifiée pour tout entier naturel non nul strictement plus petit que n avec $n \in \mathbb{N}\backslash\{0\}$ fixé et considérons l'endomorphisme orthogonal f de l'espace euclidien E de dimension n.

On commence par traiter le cas où f possède une valeur propre réelle $\lambda \in \mathbb{R}$ (i.e. le polynôme caractéristique de f possède une racine dans $\mathbb{R}$). Soit alors v un vecteur propre de f pour la valeur propre λ. On a $\|f(v)\| = \|\lambda v\| = |\lambda|\|v\|$ et, d'autre part, $\|f(v)\| = \|v\|$ car f est orthogonal. Ainsi $|\lambda|\|v\| = \|v\|$ et donc $|\lambda| = 1$ (car v est un vecteur propre de f donc $v \neq 0_E$ donc $\|v\| \neq 0$). Ainsi, $\lambda \in \{+1; -1\}$. De plus, comme $F := \mathrm{Vect}\{v\}$ est stable par f (car v est un vecteur propre de f), l'orthogonal $F^\perp$ de F est également stable par f par le lemme 2.9.3 : comme $\dim\left(F^\perp\right) = n - 1 < n$, on peut alors appliquer l'hypothèse de récurrence à l'endomorphisme orthogonal $f_{|F^\perp} : \begin{array}{ccc} F^\perp & \to & F^\perp \\ u & \mapsto & f(u) \end{array}$ de $F^\perp$ et obtenir l'existence d'une base orthonormale $\mathcal{B}_0$ de $F^\perp$ telle que

$$\mathrm{Mat}_{\mathcal{B}_0}\left(f_{|F^\perp}\right) = \begin{pmatrix} \epsilon_1 & & & & & 0 \\ & \ddots & & & & \\ & & \epsilon_r & & & \\ & & & R(\theta_1) & & \\ & & & & \ddots & \\ 0 & & & & & R(\theta_s) \end{pmatrix}$$

où $r, s \in \mathbb{N}$, pour tout $i \in \{1, \ldots, r\}$, $\epsilon_i \in \{+1; -1\}$ et, pour tout $j \in \{1, \ldots, s\}$, $R(\theta_j) := \begin{pmatrix} \cos\theta_j & -\sin\theta_j \\ \sin\theta_j & \cos\theta_j \end{pmatrix}$ avec $\theta_j \in \mathbb{R}\backslash\{k\pi \mid k \in \mathbb{Z}\}$. Considérant l'égalité $F \oplus F^\perp = E$ (proposition 2.4.5) et la base orthonormale $\mathcal{B}' := \left\{\frac{v}{\|v\|}\right\}$ de F, la famille $\mathcal{B} := \{B', \mathcal{B}_0\}$ est une base orthonormale de E et on a

$$\mathrm{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} \lambda & & & & & & 0 \\ & \epsilon_1 & & & & & \\ & & \ddots & & & & \\ & & & \epsilon_r & & & \\ & & & & R(\theta_1) & & \\ & & & & & \ddots & \\ 0 & & & & & & R(\theta_s) \end{pmatrix}$$

Supposons à présent que f ne possède pas de valeur propre réelle (i.e. le polynôme caractéristique de f ne possède pas de racine dans $\mathbb{R}$). On considère alors l'endomorphisme $h := f + f^{-1}$ de E. h est symétrique : si $u_1, u_2 \in E$,

$$\begin{aligned} \langle h(u_1), u_2\rangle &= \left\langle \left(f + f^{-1}\right)(u_1), u_2\right\rangle \\ &= \langle f(u_1), u_2\rangle + \left\langle f^{-1}(u_1), u_2\right\rangle \\ &= \left\langle f(u_1), f\left(f^{-1}(u_2)\right)\right\rangle + \left\langle f^{-1}(u_1), f^{-1}\left(f(u_2)\right)\right\rangle \\ &= \left\langle u_1, f^{-1}(u_2)\right\rangle + \langle u_1, f(u_2)\rangle \text{ (f et f^{-1} sont orthogonaux)} \\ &= \left\langle u_1, f(u_2) + f^{-1}(u_2)\right\rangle \\ &= \langle u_1, h(u_2)\rangle. \end{aligned}$$

Considérons alors une valeur propre $\lambda \in \mathbb{R}$ de h (tout endomorphisme symétrique est diagonalisable : cf théorème 2.8.4) et un vecteur propre $v \in E$ associé. D'une part, la famille $\{v, f(v)\}$ est libre, car

- $v \neq 0_E$ (car v est un vecteur propre de h),
- $f(v) \neq 0_E$ (car f est injectif puisque f est orthogonal),
- $f(v)$ ne peut s'écrire μv avec $\mu \in \mathbb{R}$ car f n'admet pas de valeur propre réelle.

D'autre part, le sev $F := \text{Vect}\{v, f(v)\}$ de E de dimension 2 est stable par f : on a $h(v) = \lambda v$ i.e. $\left(f + f^{-1}\right)(v) = \lambda v$, donc

$$\lambda f(v) = f(\lambda v) = f \circ \left(f + f^{-1}\right)(v) = f^2(v) + f \circ f^{-1}(v) = f^2(v) + v$$

d'où $f^2(v) = \lambda f(v) - v \in F$.

Considérons alors une base orthonormale $\mathcal{B}'$ de F. La matrice représentative de l'endomorphisme orthogonal $f_{|F}$ de F dans la base orthonormale $\mathcal{B}'$ est une matrice orthogonale de $\mathrm{O}_2(\mathbb{R})$ (proposition 2.6.8). $\text{Mat}_{\mathcal{B}'}(f)$ est donc de la forme $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ ou $\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$ d'après le lemme 2.9.4. Mais, comme f n'admet pas de valeur propre réelle, $\text{Mat}_{\mathcal{B}'}(f)$ ne peut pas être symétrique donc ne peut pas être de la seconde forme. Ainsi, il existe $\theta \in \mathbb{R}\backslash\{k\pi \mid k \in \mathbb{Z}\}$ tel que

$$\text{Mat}_{\mathcal{B}'}(f) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = R(\theta),$$

($\theta \notin \{k\pi \mid k \in \mathbb{Z}\}$ car $\text{Mat}_{\mathcal{B}'}(f)$ n'est pas symétrique).

Enfin, $F^\perp$ est également stable par f par le lemme 2.9.3 donc, par hypothèse de récurrence ($\dim\left(F^\perp\right) < n$), il existe une base orthonormale $\mathcal{B}_0$ de $F^\perp$ telle que

$$\text{Mat}_{\mathcal{B}'}\left(f_{|F^\perp}\right) = \begin{pmatrix} R(\theta_1) & & 0 \\ & \ddots & \\ 0 & & R(\theta_s) \end{pmatrix}$$

où, pour tout $j \in \{1, \dots, s\}$, $R(\theta_j) := \begin{pmatrix} \cos\theta_j & -\sin\theta_j \\ \sin\theta_j & \cos\theta_j \end{pmatrix}$ avec $\theta_j \in \mathbb{R}\backslash\{k\pi \mid k \in \mathbb{Z}\}$ (f ne possède pas de valeur propre réelle). En notant $\mathcal{B} := \{\mathcal{B}', \mathcal{B}_0\}$, la famille $\mathcal{B}$ est alors une base orthonormale de E et on a

$$\text{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} R(\theta) & & & 0 \\ & R(\theta_1) & & \\ & & \ddots & \\ 0 & & & R(\theta_s) \end{pmatrix}.$$

□

Remarque 2.9.5.
- Au cours de la preuve, on a montré en particulier que les seules valeurs propres réelles possibles pour un endomorphisme orthogonal sont 1 et -1.

- Si, avec les notations ci-dessus, l'endomorphisme orthogonal f possède 1 et -1 comme valeurs propres, alors les sous-espaces propres associés E_1 et E_{-1} sont orthogonaux. En effet, si $v \in E_1\backslash\{0_E\}$ et $w \in E_{-1}\backslash\{0_E\}$, on a d'une part
$$\langle f(v), f(w)\rangle = \langle v, -w\rangle = -\langle v, w\rangle,$$
et d'autre part
$$\langle f(v), f(w)\rangle = \langle v, w\rangle$$
(car f est orthogonal), ainsi $\langle v, w\rangle = -\langle v, w\rangle$ donc $\langle v, w\rangle = 0$.

  Par ailleurs, le sous-espace vectoriel $E_1 \oplus E_{-1}$ de E est stable par f donc $(E_1 \oplus E_{-1})^{\perp}$ également.

  Au total, pour calculer une réduction de f comme dans le théorème 2.9.1, on peut donc commencer par déterminer une base orthonormale $\mathcal{B}'$ de E_1, une base orthonormale $\mathcal{B}''$ de E_{-1} puis une base orthonormale du sous-espace stable $\{\mathcal{B}', \mathcal{B}''\}^{\perp}$. Dans la représentation matricielle correspondante de f, le bloc correspondant à la restriction de f à ce dernier sous-espace stable est orthogonal sans valeur propre réelle et on peut alors lui appliquer la méthode de la preuve du théorème pour ce cas.

Exemple 2.9.6. Considérons la matrice orthogonale
$$A := \frac{1}{3}\begin{pmatrix} 2 & -1 & 2 \\ 2 & 2 & -1 \\ -1 & 2 & 2 \end{pmatrix}$$
de l'exemple 2.6.11 1. Calculons son polynôme caractéristique : on a
$$\begin{aligned}
\chi_A = \det(A - XI_3) &= \begin{vmatrix} \frac{2}{3} - X & -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} - X & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} - X \end{vmatrix} \\
&\underset{C_1 \leftarrow C_1 + C_2 + C_3}{=} \begin{vmatrix} 1 - X & -\frac{1}{3} & \frac{2}{3} \\ 1 - X & \frac{2}{3} - X & -\frac{1}{3} \\ 1 - X & \frac{2}{3} & \frac{2}{3} - X \end{vmatrix} \\
&= (1 - X)\begin{vmatrix} 1 & -\frac{1}{3} & \frac{2}{3} \\ 1 & \frac{2}{3} - X & -\frac{1}{3} \\ 1 & \frac{2}{3} & \frac{2}{3} - X \end{vmatrix} \\
&\underset{L_2 \leftarrow L_2 - L_1,\ L_3 \leftarrow L_3 - L_1}{=} (1 - X)\begin{vmatrix} 1 & -\frac{1}{3} & \frac{2}{3} \\ 0 & 1 - X & -1 \\ 0 & 1 & -X \end{vmatrix} \\
&= (1 - X)\left(X^2 - X + 1\right)
\end{aligned}$$

On a $E_1 = \operatorname{Ker}\left(\frac{1}{3}\begin{pmatrix} -1 & -1 & 2 \\ 2 & -1 & -1 \\ -1 & 2 & -1 \end{pmatrix}\right)$ et le vecteur colonne $Y_1 := \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ de norme 1 engendre E_1.

De plus, $\left(\operatorname{Vect}\left\{\frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix}\right\}\right)^{\perp} = \left\{\begin{pmatrix}x\\y\\z\end{pmatrix} \in \mathrm{M}_{3,1}(\mathbb{R}) \mid x+y+z=0\right\}$, dont une base orthonormale est formée des vecteurs $Y_2 := \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\\0\end{pmatrix}$ et $Y_3 := \frac{1}{\sqrt{6}}\begin{pmatrix}1\\1\\-2\end{pmatrix}$.

On a

$$AY_2 = A\begin{pmatrix}\frac{1}{\sqrt{2}}\\-\frac{1}{\sqrt{2}}\\0\end{pmatrix} = \begin{pmatrix}\frac{1}{\sqrt{2}}\\0\\-\frac{1}{\sqrt{2}}\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\-1\end{pmatrix} = \frac{1}{2\sqrt{2}}\left(\sqrt{2}Y_2+\sqrt{6}Y_3\right) = \frac{1}{2}Y_2 + \frac{\sqrt{3}}{2}Y_3$$

et

$$AY_3 = A\begin{pmatrix}\frac{1}{\sqrt{6}}\\\frac{1}{\sqrt{6}}\\-\frac{2}{\sqrt{6}}\end{pmatrix} = \begin{pmatrix}-\frac{1}{\sqrt{6}}\\\frac{2}{\sqrt{6}}\\-\frac{1}{\sqrt{6}}\end{pmatrix} = \frac{1}{\sqrt{6}}\begin{pmatrix}-1\\2\\-1\end{pmatrix} = \frac{1}{2\sqrt{6}}\left(-3\sqrt{2}Y_2+\sqrt{6}Y_3\right) = -\frac{\sqrt{3}}{2}Y_2 + \frac{1}{2}Y_3.$$

Ainsi, si on note

$$P := \begin{pmatrix}\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}}\\\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}}\\\frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}}\end{pmatrix} \in \mathrm{O}_3(\mathbb{R}),$$

on a

$$P^{-1}AP = {}^tPAP = \begin{pmatrix}1 & 0 & 0\\0 & \frac{1}{2} & -\frac{\sqrt{3}}{2}\\0 & \frac{\sqrt{3}}{2} & \frac{1}{2}\end{pmatrix} = \begin{pmatrix}1 & 0 & 0\\0 & \cos\left(\frac{\pi}{3}\right) & -\sin\left(\frac{\pi}{3}\right)\\0 & \sin\left(\frac{\pi}{3}\right) & \cos\left(\frac{\pi}{3}\right)\end{pmatrix}$$

La matrice A est donc la matrice représentative dans la base canonique de $\mathbb{R}^3$ de la rotation d'axe la droite vectorielle engendrée par le vecteur $(1,1,1)$ et d'angle $\frac{\pi}{3}$.

Chapitre 3

Espaces hermitiens

3.1 Introduction

On étudie l'analogue de la notion de produit scalaire pour les espaces vectoriels sur $\mathbb{C}$: le produit scalaire hermitien. La plupart des notions introduites et des résultats énoncés dans le chapitre précédent auront leurs analogues dans le cadre hermitien. Nous verrons en particulier que tout endomorphisme unitaire (version hermitienne de la notion d'endomorphisme orthogonal) est diagonalisable dans une base orthonormale. Nous remarquerons également les différences qui existent entre les espaces hermitiens et les espaces euclidiens.

3.2 Produit scalaire hermitien sur un espace vectoriel complexe

Soit E un espace vectoriel sur $\mathbb{C}$.

Définition 3.2.1. *Considérons une application* $\langle \cdot,\cdot \rangle : \begin{array}{ccc} E \times E & \to & \mathbb{C} \\ (v,w) & \mapsto & \langle v,w \rangle \end{array}$. *On dit que* $\langle \cdot,\cdot \rangle$ *est un produit scalaire hermitien sur E si* $\langle \cdot,\cdot \rangle$ *est*

1. *sesquilinéaire (à gauche), i.e. pour tous* $v, w, v_1, v_2, w_1, w_2 \in E$ *et tous* $\lambda, \mu \in \mathbb{C}$, $\langle \lambda v_1 + \mu v_2, w\rangle = \overline{\lambda}\langle v_1, w\rangle + \overline{\mu}\langle v_2, w\rangle$ *et* $\langle v, \lambda w_1 + \mu w_2\rangle = \lambda\langle v, w_1\rangle + \mu\langle v, w_2\rangle$,
2. *symétrique hermitienne, i.e. pour tous* $v, w \in E$, $\langle w, v\rangle = \overline{\langle v, w\rangle}$,
3. *définie positive, i.e. pour tout* $v \in E$, $\langle v, v\rangle$ *est un réel positif ou nul et* $\langle v, v\rangle = 0$ *si et seulement si* $v = 0_E$.

Exemple 3.2.2. 1. Soit $n \in \mathbb{N}\backslash\{0\}$, pour tous $v = (x_1, \dots, x_n)$ et $w = (y_1, \dots, y_n)$ dans $\mathbb{C}^n$, on définit

$$\langle v, w\rangle_{\text{can}} := \overline{x_1}y_1 + \cdots + \overline{x_n}y_n.$$

L'application $\langle \cdot,\cdot \rangle_{\text{can}} : \begin{array}{ccc} \mathbb{C}^n \times \mathbb{C}^n & \to & \mathbb{C} \\ (v,w) & \mapsto & \langle v,w \rangle_{\text{can}} \end{array}$ est alors un produit scalaire hermitien sur $\mathbb{C}^n$, appelé produit scalaire hermitien canonique sur $\mathbb{C}^n$. En effet, soient $v = (x_1, \dots, x_n), v' = (x'_1, \dots, x'_n), w = (y_1, \dots, y_n), w' = (y'_1, \dots, y'_n) \in \mathbb{C}^n$, soient $\lambda, \mu \in \mathbb{C}$, on a

- $$\begin{aligned}\langle \lambda v + \mu v', w\rangle_{\mathrm{can}} &= \sum_{k=1}^{n} \overline{(\lambda x_k + \mu x'_k)} y_k \\ &= \overline{\lambda} \sum_{k=1}^{n} \overline{x_k} y_k + \overline{\mu} \sum_{k=1}^{n} \overline{x'_k} y_k \\ &= \overline{\lambda}\langle v, w\rangle_{\mathrm{can}} + \overline{\mu}\langle v', w\rangle_{\mathrm{can}},\end{aligned}$$

- $$\begin{aligned}\langle v, \lambda w + \mu w'\rangle_{\mathrm{can}} &= \sum_{k=1}^{n} \overline{x_k} \left(\lambda y_k + \mu y'_k\right) \\ &= \lambda \sum_{k=1}^{n} \overline{x_k} y_k + \mu \sum_{k=1}^{n} \overline{x_k} y'_k \\ &= \lambda\langle v, w\rangle_{\mathrm{can}} + \mu\langle v, w'\rangle_{\mathrm{can}},\end{aligned}$$

- $$\begin{aligned}\langle w, v\rangle_{\mathrm{can}} &= \sum_{k=1}^{n} \overline{y_k} x_k \\ &= \sum_{k=1}^{n} x_k \overline{y_k} \\ &= \overline{\sum_{k=1}^{n} \overline{x_k} y_k} \\ &= \overline{\langle w, v\rangle_{\mathrm{can}}},\end{aligned}$$

- $$\langle v, v\rangle_{\mathrm{can}} = \sum_{k=1}^{n} \overline{x_k} x_k = \sum_{k=1}^{n} |x_k|^2 \geqslant 0$$

  et $\langle v, v\rangle_{\mathrm{can}} = \sum_{k=1}^{n} |x_k|^2 = 0$ ssi pour tout $k \in \{1, \ldots, n\}$, $|x_k| = 0$ ssi pour tout $k \in \{1, \ldots, n\}$, $x_k = 0$ ssi $v = (0, \ldots, 0)$.

2. Supposons que le $\mathbb{C}$-espace vectoriel E soit de dimension finie $n \in \mathbb{N}\backslash\{0\}$ et soit $\mathcal{B} = \{e_1, \ldots, e_n\}$ une base de E. Pour tous vecteurs v et w de E, de coordonnées respectives $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ dans la base $\mathcal{B}$, on définit

$$\langle v, w\rangle_{\mathcal{B}} := \overline{x_1} y_1 + \cdots + \overline{x_n} y_n = {}^t\overline{\mathrm{Mat}_{\mathcal{B}}(v)}\mathrm{Mat}_{\mathcal{B}}(w).$$

L'application $\langle\cdot,\cdot\rangle_{\mathcal{B}} : \begin{array}{ccc} E\times E & \to & \mathbb{C} \\ (v,w) & \mapsto & \langle v,w\rangle_{\mathcal{B}} \end{array}$ est alors un produit scalaire hermitien sur E, appelé produit scalaire hermitien associé à la base $\mathcal{B}$.

3. Soit $n\in\mathbb{N}\backslash\{0\}$. Pour toutes matrices $A=(a_{i\,j})_{1\leqslant i,j\leqslant n}$ et $B=(b_{i\,j})_{1\leqslant i,j\leqslant n}$ dans $\mathrm{M}_n(\mathbb{C})$, on définit
$$\langle A,B\rangle := \sum_{1\leqslant i,j\leqslant n} \overline{a_{i\,j}}b_{i\,j} = \mathrm{Tr}\left({}^t\overline{A}\,B\right).$$
L'application $\langle\cdot,\cdot\rangle : \begin{array}{ccc} \mathrm{M}_n(\mathbb{C})\times\mathrm{M}_n(\mathbb{C}) & \to & \mathbb{C} \\ (A,B) & \mapsto & \langle A,B\rangle \end{array}$ est alors un produit scalaire hermitien sur $\mathrm{M}_n(\mathbb{C})$: il s'agit du produit scalaire hermitien associé à la base canonique $\{E_{i\,j}\}_{1\leqslant i,j\leqslant n}$ de $\mathrm{M}_n(\mathbb{C})$.

4. Soit $n\in\mathbb{N}\backslash\{0\}$. Pour tous polynômes P et Q de $\mathbb{C}_n[X]$, on définit
$$\langle P,Q\rangle := \int_0^1 \overline{P(t)}Q(t)\mathrm{d}t.$$
L'application $\langle\cdot,\cdot\rangle : \begin{array}{ccc} \mathbb{C}_n[X]\times\mathbb{C}_n[X] & \to & \mathbb{C} \\ (P,Q) & \mapsto & \langle P,Q\rangle \end{array}$ est alors un produit scalaire hermitien sur $\mathbb{C}_n[X]$. La sesquilinéarité et la symétrie hermitienne de l'application $\langle\cdot,\cdot\rangle$ proviennent de la linéarité de l'intégrale et du fait que, pour toute fonction continue $f : [0;1]\to\mathbb{C}$, $\overline{\int_0^1 f(t)\mathrm{d}t} = \int_0^1 \overline{f}(t)\mathrm{d}t$. Montrons que $\langle\cdot,\cdot\rangle$ est bien définie positive. Soit $P\in\mathbb{C}_n[X]$, on a $\langle P,P\rangle = \int_0^1 |P(t)|^2\mathrm{d}t \geqslant 0$ (l'intégrale sur le segment $[0;1]$ de la fonction à valeurs réelles $\mathbb{R}\to\mathbb{R}\ ;\ t\mapsto|P(t)|^2$ continue positive est positive). De plus, si $\langle P,P\rangle = \int_0^1 |P(t)|^2\mathrm{d}t = 0$, comme la fonction à valeurs réelles $\mathbb{R}\to\mathbb{R}\ ;\ t\mapsto|P(t)|^2$ est continue et positive, on a, pour tout $t\in[0,1]$, $|P(t)|^2=0$ (l'intégrale d'une fonction réelle continue et positive sur un segment est nulle si et seulement si la fonction est identiquement nulle sur ce segment) et donc, pour tout $t\in[0,1]$, $P(t)=0$, donc le polynôme P est nul (un polynôme ayant une infinité de racines est nécessairement le polynôme nul).

Remarque 3.2.3. Soit $n\in\mathbb{N}\backslash\{0\}$. La forme bilinéaire symétrique
$$b(\cdot,\cdot) : \begin{array}{ccc} \mathbb{C}^n\times\mathbb{C}^n & \to & \mathbb{C} \\ \big((x_1,\ldots,x_n),(y_1,\ldots,y_n)\big) & \mapsto & \sum_{k=1}^n x_k y_k \end{array}$$
n'est

- ni sesquilinéaire, car, par exemple,
$$b\big(i\cdot(1,0,\ldots,0),(1,0,\ldots,0)\big) = b\big((i,0,\ldots,0),(1,0,\ldots,0)\big) = i \quad\text{et}\quad \bar{i}b\big((1,0,\ldots,0),(1,0,\ldots,0)\big) = -i,$$
- ni symétrique hermitienne, car, par exemple,
$$b\big((i,0,\ldots,0),(1,0,\ldots,0)\big) = i \quad\text{et}\quad \overline{b\big((1,0,\ldots,0),(i,0,\ldots,0)\big)} = -i,$$

- ni positive, car, par exemple, $b\big((1+i,0,\dots,0),(1+i,0,\dots,0)\big) = (1+i)^2 = 2i$,
- ni définie si $n \geqslant 2$, car, par exemple, $b\big((1,i,0,\dots,0),(1,i,0,\dots,0)\big) = 0$.

Définition 3.2.4. *Si E est un espace vectoriel de dimension finie muni d'un produit scalaire hermitien $\langle\cdot,\cdot\rangle$, le couple $(E,\langle\cdot,\cdot\rangle)$ est appelé espace hermitien.*

La plupart des propriétés que nous avons montrées pour les espaces euclidiens vont avoir leurs analogues pour les espaces hermitiens.

Dans la suite de cette section, on suppose que E est un $\mathbb{C}$-espace vectoriel de dimension finie muni d'un produit scalaire hermitien $\langle\cdot,\cdot\rangle$. En premier lieu, si l'on note, pour tout vecteur v de E, $\|v\| := \sqrt{\langle v,v\rangle}$ ($\langle v,v\rangle$ est un nombre réel positif), l'inégalité de Cauchy-Schwarz est vérifiée et l'application $\|\cdot\| : \begin{array}{ccc} E & \to & [0,+\infty[\\ v & \mapsto & \|v\| \end{array}$ est une norme sur E :

Lemme 3.2.5 (Inégalité de Cauchy-Schwarz)**.** *Pour tous vecteurs v et w de E, on a l'inégalité*

$$|\langle v,w\rangle| \leqslant \|v\|\|w\|,$$

et $|\langle v,w\rangle| = \|v\|\|w\|$ si et seulement si les vecteurs v et w sont liés.

Démonstration. Soient $v,w \in E$. Si $\langle v,w\rangle = 0$, l'inégalité de Cauchy-Schwarz est vérifiée (car on a toujours $\|v\|\|w\| \geqslant 0$). Supposons donc que $\langle v,w\rangle \neq 0$ (en particulier, $v \neq 0_E$ et $w \neq 0_E$). On pose ensuite $\alpha := \frac{1}{\langle v,w\rangle}$ et $\widetilde{v} := \alpha v$. Alors $\langle \widetilde{v},w\rangle = \langle \alpha v,w\rangle = \overline{\alpha}\langle v,w\rangle = 1$ et on a, pour tout $\lambda \in \mathbb{R}$,

$$\begin{aligned} 0 \leqslant \|\widetilde{v}+\lambda w\|^2 &= \langle \widetilde{v}+\lambda w, \widetilde{v}+\lambda w\rangle \\ &= \langle \widetilde{v},\widetilde{v}\rangle + \lambda\langle \widetilde{v},w\rangle + \lambda\langle w,\widetilde{v}\rangle + \lambda^2\langle w,w\rangle \ (\lambda \in \mathbb{R} \text{ donc } \overline{\lambda} = \lambda) \\ &= \|\widetilde{v}\|^2 + 2\lambda + \lambda^2\|w\|^2 \ (\langle w,\widetilde{v}\rangle = \overline{\langle \widetilde{v},w\rangle} = \overline{1} = 1). \end{aligned}$$

La fonction polynomiale du second degré

$$p : \begin{array}{ccc} \mathbb{R} & \to & \mathbb{R} \\ \lambda & \mapsto & \|w\|^2\lambda^2 + 2\lambda + \|\widetilde{v}\|^2 \end{array}$$

étant positive sur tout $\mathbb{R}$, le discriminant associé $4-4\|w\|^2\,\|\widetilde{v}\|^2$ est négatif ou nul i.e. $1 \leqslant \|\widetilde{v}\|\,\|w\|$. Or

$$\|\widetilde{v}\|\,\|w\| = \|\alpha v\|\|w\| = |\alpha|\|v\|\|w\| = \frac{1}{|\langle v,w\rangle|}\|v\|\|w\|.$$

Ainsi, on a bien $|\langle v,w\rangle| \leqslant \|v\|\|w\|$.

Traitons maintenant le cas d'égalité de l'inégalité de Cauchy-Schwarz. Supposons donc que $|\langle v,w\rangle| = \|v\|\|w\|$. Alors,

- si $\langle v,w\rangle = 0$: on a $|\langle v,w\rangle| = \|v\|\|w\| \Leftrightarrow 0 = \|v\|\|w\| \Leftrightarrow \|v\| = 0$ ou $\|w\| = 0 \Leftrightarrow v = 0_E$ ou $w = 0_E$, et les vecteurs v et w sont en particulier liés,

- si $\langle v, w\rangle \neq 0$: on a $|\langle v, w\rangle| = \|v\|\|w\| \Leftrightarrow 1 = \|w\|^2 \|\widetilde{v}\|^2$ et le polynôme associé à la fonction polynomiale du second degré p (avec les mêmes notations que ci-dessus) possède une racine (double) $\lambda_0 \in \mathbb{R}$ et donc $\|\widetilde{v} + \lambda_0 w\|^2 = 0$ i.e. $\alpha v + \lambda_0 w = 0_E$: les vecteurs v et w sont donc liés (α est non nul).

□

Corollaire 3.2.6. *Le couple $(E, \|\cdot\|)$ est un espace vectoriel normé.*

Démonstration. La preuve est analogue à la preuve du corollaire 2.2.7 : si $v \in E$ et $\lambda \in \mathbb{C}$, on a

$$\begin{aligned} \|\lambda v\| &= \sqrt{\langle \lambda v, \lambda v\rangle} \\ &= \sqrt{\overline{\lambda}\lambda\langle v, v\rangle} \\ &= \sqrt{|\lambda|^2}\sqrt{\langle v, v\rangle} \\ &= |\lambda|\|v\| \end{aligned}$$

et $\|v\| = \sqrt{\langle v, v\rangle} = 0$ ssi $\langle v, v\rangle = 0$ ssi $v = 0_E$.

Enfin, si $v, w \in E$,

$$\begin{aligned} \|v+w\|^2 &= \langle v+w, v+w\rangle \\ &= \langle v, v\rangle + \langle v, w\rangle + \langle w, v\rangle + \langle w, w\rangle \\ &= \langle v, v\rangle + \langle v, w\rangle + \overline{\langle v, w\rangle} + \langle w, w\rangle \\ &= \|v\|^2 + 2\operatorname{Re}(\langle v, w\rangle) + \|w\|^2 \\ &\leqslant \|v\|^2 + 2|\langle v, w\rangle| + \|w\|^2 \text{ (si } z \in \mathbb{C},\ \operatorname{Re}(z) \leqslant |z|) \\ &\leqslant \|v\|^2 + 2\|v\|^2\|w\|^2 + \|w\|^2 \text{ (par l'inégalité de Cauchy-Schwarz)} \\ &= (\|v\| + \|w\|)^2 . \end{aligned}$$

□

La norme $\|\cdot\|$ est appelée norme hermitienne associée au produit scalaire hermitien $\langle\cdot,\cdot\rangle$.

Exemple 3.2.7. Soit $n \in \mathbb{N}\backslash\{0\}$. Sur $\mathbb{C}^n$, la norme hermitienne associée à au produit scalaire hermitien canonique vérifie, pour tout $v = (x_1, \ldots, x_n) \in \mathbb{C}^n$,

$$\|v\| = \sqrt{\langle v, v\rangle_{\text{can}}} = \sqrt{\sum_{k=1}^{n} \overline{x_k} x_k} = \sqrt{\sum_{k=1}^{n} |x_k|^2}.$$

L'analogue hermitien de l'expression du produit scalaire à l'aide de la norme euclidienne (identité dite de polarisation, énoncée dans la remarque 2.2.8) est l'identité suivante :

Lemme 3.2.8. *Pour tous $v, w \in E$,*

$$\langle v, w\rangle = \frac{1}{2}\left(\|v+w\|^2 + i\|iv+w\|^2 - (1+i)\|v\|^2 - (1+i)\|w\|^2\right).$$

Démonstration. Soient $v, w \in E$. D'après les dernières égalités de la preuve du corollaire 3.2.6, on a
$$\mathrm{Re}(\langle v, w\rangle) = \frac{1}{2}\left(\|v+w\|^2 - \|v\|^2 - \|w\|^2\right).$$
On a donc également
$$\mathrm{Re}(\langle iv, w\rangle) = \frac{1}{2}\left(\|iv+w\|^2 - \|iv\|^2 - \|w\|^2\right) = \frac{1}{2}\left(\|iv+w\|^2 - \|v\|^2 - \|w\|^2\right).$$
Or $\mathrm{Re}(\langle iv, w\rangle) = \mathrm{Re}(\bar{i}\langle v, w\rangle) = \mathrm{Re}(-i\langle v, w\rangle) = \mathrm{Im}(\langle v, w\rangle)$ donc
$$\mathrm{Im}(\langle v, w\rangle) = \frac{1}{2}\left(\|iv+w\|^2 - \|v\|^2 - \|w\|^2\right).$$
Au total,
$$\begin{aligned}
\langle v, w\rangle &= \mathrm{Re}(\langle v, w\rangle) + i\,\mathrm{Im}(\langle v, w\rangle)\\
&= \frac{1}{2}\left(\|v+w\|^2 - \|v\|^2 - \|w\|^2\right) + \frac{i}{2}\left(\|iv+w\|^2 - \|v\|^2 - \|w\|^2\right)\\
&= \frac{1}{2}\left(\|v+w\|^2 + i\|iv+w\|^2 - (1+i)\|v\|^2 - (1+i)\|w\|^2\right).
\end{aligned}$$
□

3.3 Orthogonalité dans les espaces hermitiens

Soit $(E, \langle\cdot,\cdot\rangle)$ un espace hermitien.

À l'instar du produit scalaire euclidien, le produit scalaire hermitien nous permet de définir une notion d'orthogonalité :

Définition 3.3.1. *Soient v et w deux vecteurs de E. On dit que v et w sont orthogonaux si $\langle v, w\rangle = 0$. Dans ce cas, on dira aussi que v est orthogonal à w et que w est orthogonal à v ($\langle v, w\rangle = 0$ si et seulement si $\langle w, v\rangle = 0$ car $\langle w, v\rangle = \overline{\langle v, w\rangle}$).*

Exemple 3.3.2. Dans $\mathbb{C}^2$ muni du produit scalaire hermitien canonique, les vecteurs $(i, 1)$ et $(i, -1)$ sont orthogonaux : $\langle (i,1), (i,-1)\rangle_{\mathrm{can}} = \bar{i} \times i + \bar{1} \times (-1) = 1 - 1 = 0$.

La version du théorème de Pythagore adaptée au cadre hermitien est l'énoncé suivant :

Lemme 3.3.3 (Théorème de Pythagore)**.** *Soient v et w deux vecteurs de E. Si v et w sont orthogonaux, alors $\|v+w\|^2 = \|v\|^2 + \|w\|^2$.*

Démonstration. On a $\|v+w\|^2 = \|v\|^2 + 2\,\mathrm{Re}(\langle v, w\rangle) + \|w\|^2$. Ainsi, si v et w sont orthogonaux i.e. $\langle v, w\rangle = 0$, on a bien $\|v+w\|^2 = \|v\|^2 + \|w\|^2$. □

Remarque 3.3.4. Dans le cadre hermitien, la réciproque du théorème de Pythagore n'est plus vraie en général. Par exemple, si on considère les vecteurs $v := (1, 0)$ et $w := (i, 0)$ de $\mathbb{C}^2$ muni du produit scalaire hermitien canonique, on a $\|v\|^2 = 1$, $\|w\|^2 = 1$ et $\|v+w\|^2 = \|(1+i, 0)\|^2 = 2 = \|v\|^2 + \|w\|^2$, alors que les v et w ne sont pas orthogonaux.

Une équivalence vraie est la suivante : avec les notations du lemme 3.3.3, v et w sont orthogonaux si et seulement $\|v+w\|^2 = \|v\|^2 + \|w\|^2$ et $\|iv+w\|^2 = \|v\|^2 + \|w\|^2$. En effet, si ces deux dernières égalités sont satisfaites alors, d'après les relations établies dans la preuve du lemme 3.2.8, $\mathrm{Re}(\langle v, w\rangle) = 0$ et $\mathrm{Im}(\langle v, w\rangle) = 0$ i.e. $\langle v, w\rangle = 0$.

Les notions de base orthogonale et de base orthonormale pour un espace hermitien sont exactement les mêmes que dans le cadre euclidien :

Définition 3.3.5. *Soit $\mathcal{B} = \{e_1, \ldots, e_n\}$ une base de E. On dit que $\mathcal{B}$ est*

- *une base orthogonale de E si pour tous $k, l \in \{1, \ldots, n\}$ tels que $k \neq l$, $\langle e_k, e_l\rangle = 0$,*
- *une base orthonormale de E si pour tous $k, l \in \{1, \ldots, n\}$, $\langle e_k, e_l\rangle = \delta_{k\,l}$.*

Exemple 3.3.6.
- Si $n \in \mathbb{N}\backslash\{0\}$, la base canonique de $\mathbb{C}^n$ est une base orthonormale pour le produit scalaire hermitien canonique sur $\mathbb{C}^n$ (exemple 3.2.2 1.).
- Une base $\mathcal{B}$ de E est une base orthonormale pour le produit scalaire hermitien associé $\langle\cdot,\cdot\rangle_{\mathcal{B}}$ (exemple 3.2.2 2.).

De façon analogue au cadre euclidien (cf. remarque 2.3.6), si $\mathcal{B} = \{e_1, \ldots, e_n\}$ est une base orthonormale pour $\langle\cdot,\cdot\rangle$ et $v, w \in E$, on a

$$v = \sum_{k=1}^{n} \overline{\langle v, e_k\rangle}\, e_k = \sum_{k=1}^{n} \langle e_k, v\rangle\, e_k \text{ et } \langle v, w\rangle = \sum_{k=1}^{n} \langle v, e_k\rangle \overline{\langle w, e_k\rangle} = \sum_{k=1}^{n} \langle v, e_k\rangle\langle e_k, w\rangle.$$

On définit également les notions de famille orthogonale de E et de famille orthonormale de E (de telles familles sont des familles libres de E) et on a aussi un procédé d'orthonormalisation de Gram-Schmidt dans le cadre hermitien :

Théorème 3.3.7 (Procédé d'orthonormalisation de Gram-Schmidt)**.** *Soit $\{v_1, \ldots, v_p\}$ une famille libre de E. On peut construire, de façon algorithmique, une famille orthonormale $\{e_1, \ldots, e_p\}$ de E telle que, pour tout $k \in \{1, \ldots, p\}$, $\mathrm{Vect}\{e_1, \ldots, e_k\} = \mathrm{Vect}\{v_1, \ldots, v_k\}$ (en particulier, $\{e_1, \ldots, e_p\}$ est une base orthonormale de $\mathrm{Vect}\{v_1, \ldots, v_p\}$).*

Démonstration. Comme dans la preuve du théorème 2.3.7, on construit, de façon récursive, une base orthogonale $\{\epsilon_1, \ldots, \epsilon_p\}$ pour $\mathrm{Vect}\{v_1, \ldots, v_p\}$. On en déduit une base orthonormale en "normalisant" les vecteurs obtenus.

Le procédé récursif est le suivant : pour $1 \leqslant k \leqslant p-1$, on suppose que l'on a déjà construit des vecteurs $\epsilon_1, \ldots, \epsilon_k \in E$ orthogonaux deux à deux tels que $\mathrm{Vect}\{\epsilon_1, \ldots, \epsilon_k\} = \mathrm{Vect}\{v_1, \ldots, v_k\}$. On recherche alors un vecteur ϵ_{k+1} de E de la forme

$$\epsilon_{k+1} = v_{k+1} + \lambda_1\epsilon_1 + \cdots + \lambda_k\epsilon_k \in \mathrm{Vect}\{\epsilon_1, \ldots, \epsilon_k, v_{k+1}\} = \mathrm{Vect}\{v_1, \ldots, v_k, v_{k+1}\},$$

avec $\lambda_1, \ldots, \lambda_k \in \mathbb{C}$, tel que, pour tout $i \in \{1, \ldots, k\}$, $\langle\epsilon_{k+1}, \epsilon_i\rangle = 0$.

Or, pour $i \in \{1, \ldots, k\}$,

$$\begin{aligned}\langle\epsilon_{k+1}, \epsilon_i\rangle = 0 \quad &\Leftrightarrow \quad \langle v_{k+1} + \lambda_1\epsilon_1 + \cdots + \lambda_k\epsilon_k, \epsilon_i\rangle = 0\\ &\Leftrightarrow \quad \langle v_{k+1}, \epsilon_i\rangle + \overline{\lambda_i}\langle\epsilon_i, \epsilon_i\rangle = 0\\ &\Leftrightarrow \quad \lambda_i = -\frac{\overline{\langle v_{k+1}, \epsilon_i\rangle}}{\|\epsilon_i\|^2}\end{aligned}$$

Ainsi, par construction, la famille $\{\epsilon_1, \ldots, \epsilon_{k+1}\}$ avec $\epsilon_{k+1} := v_{k+1} - \sum_{i=1}^k \frac{\overline{\langle v_{k+1}, \epsilon_i \rangle}}{\|\epsilon_i\|^2} \epsilon_i$ est orthogonale et engendre bien $\mathrm{Vect}\{v_1, \ldots, v_k, v_{k+1}\}$ car $\mathrm{Vect}\{v_1, \ldots, v_k, v_{k+1}\} = \mathrm{Vect}\{\epsilon_1, \ldots, \epsilon_k, v_{k+1}\} = \mathrm{Vect}\{\epsilon_1, \ldots, \epsilon_k, \epsilon_{k+1}\}$ (car $\epsilon_{k+1} = v_{k+1} - \sum_{i=1}^k \frac{\overline{\langle v_{k+1}, \epsilon_i \rangle}}{\|\epsilon_i\|^2} \epsilon_i \in \mathrm{Vect}\{\epsilon_1, \ldots, \epsilon_k, v_{k+1}\}$ et $v_{k+1} = \epsilon_{k+1} + \sum_{i=1}^k \frac{\overline{\langle v_{k+1}, \epsilon_i \rangle}}{\|\epsilon_i\|^2} \epsilon_i \in \mathrm{Vect}\{\epsilon_1, \ldots, \epsilon_k, \epsilon_{k+1}\}$).

□

Exemple 3.3.8. On détermine une base orthonormale du sous-espace vectoriel

$$F := \left\{(x, y, z, t) \in \mathbb{C}^4 \mid x - y + iz - it = 0\right\}$$

de $\mathbb{C}^4$ par rapport au produit scalaire hermitien canonique. On commence par remarquer que les vecteurs $v_1 := (1, 1, 0, 0)$, $v_2 := (1, 0, i, 0)$ et $v_3 := (0, 0, 1, 1)$ forment une base (non orthogonale) de F. On applique ensuite le procédé d'orthonormalisation de Gram-Schmidt : on pose

$$\begin{aligned}
\epsilon_1 &:= v_1 = (1, 1, 0, 0) \\
\epsilon_2 &:= v_2 - \frac{\overline{\langle v_2, \epsilon_1 \rangle}}{\|\epsilon_1\|^2} \epsilon_1 = v_2 - \frac{1}{2}\epsilon_1 = \left(\frac{1}{2}, -\frac{1}{2}, i, 0\right) = \frac{1}{2}(1, -1, 2i, 0) \\
\epsilon_2' &:= (1, -1, 2i, 0) \\
\epsilon_3 &:= v_3 - \frac{\overline{\langle v_3, \epsilon_1 \rangle}}{\|\epsilon_1\|^2} \epsilon_1 - \frac{\overline{\langle v_3, \epsilon_2' \rangle}}{\|\epsilon_2'\|^2} \epsilon_2' = v_3 + \frac{2i}{6}\epsilon_2' = \left(\frac{i}{3}, -\frac{i}{3}, \frac{1}{3}, 1\right) = \frac{1}{3}(i, -i, 1, 3) \\
\epsilon_3' &:= (i, -i, 1, 3).
\end{aligned}$$

La famille $\{\epsilon_1, \epsilon_2', \epsilon_3'\}$ est alors une base orthogonale de F et la famille

$$\left\{\frac{1}{\sqrt{2}}(1, 1, 0, 0), \frac{1}{\sqrt{6}}(1, -1, 2i, 0), \frac{1}{\sqrt{12}}(i, -i, 1, 3)\right\}$$

est une base orthonormale de F.

En particulier, le théorème 3.3.7 permet de montrer l'existence d'une base orthonormale pour l'espace hermitien $(E, \langle \cdot, \cdot \rangle)$. Il permet également de montrer que toute famille orthonormale de E peut être complétée en une base orthonormale de E.

Si maintenant A est un sous-ensemble de E, on définit aussi l'orthogonal

$$A^\perp := \{v \in E \mid \forall w \in A, \langle v, w \rangle = 0\}$$

de A, qui est un sous-espace vectoriel de E.

Exemple 3.3.9. Dans $\mathbb{C}^3$ muni du produit scalaire hermitien canonique, on a

$$\begin{aligned}
\{(-2, 3 + 2i, -i)\}^\perp &= \left\{(x, y, z) \in \mathbb{C}^3 \mid \langle (x, y, z), (-2, 3 + 2i, -i) \rangle_{\mathrm{can}} = -2\overline{x} + (3 - 2i)\overline{y} + i\overline{z} = 0\right\} \\
&= \left\{(x, y, z) \in \mathbb{C}^3 \mid -2x + (3 + 2i)y - iz = 0\right\}.
\end{aligned}$$

Si F est un sous-espace vectoriel de E, on retrouve, par des arguments analogues au cadre euclidien (cf. proposition 2.4.5 et sa preuve), les propriétés suivantes :

Proposition 3.3.10. *On a :*

1. $E = F \oplus F^{\perp}$,
2. $\left(F^{\perp}\right)^{\perp} = F$.

On peut ainsi définir la projection orthogonale sur F, i.e. la projection sur F parallèlement à $F^{\perp}$, notée p_F, et la symétrie orthogonale par rapport à F, i.e. la symétrie par rapport à F parallèlement à $F^{\perp}$, notée s_F.

Comme dans le cadre euclidien, si $v \in E$, on peut calculer la distance hermitienne de v à F $d(v,F) := \inf_{w \in F} \|w - v\|$ à l'aide de la projection orthogonale sur F : on a

$$d(v,F) = \|v - p_F(v)\| = \left\| v - \sum_{k=1}^{p} \overline{\langle v, e_k \rangle}\, e_k \right\|,$$

si $\{e_1, \ldots, e_p\}$ est une base orthonormale de F.

3.4 Représentation matricielle du produit scalaire hermitien

Soit $(E, \langle \cdot,\cdot \rangle)$ un espace hermitien et notons $n := \dim(E)$. Si $\mathcal{B} = \{e_1, \ldots, e_n\}$ est une base de E et si v et w sont deux vecteurs de E de coordonnées respectives $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ dans la base $\mathcal{B}$, alors

$$\begin{aligned}
\langle v, w \rangle &= \left\langle \sum_{k=1}^{n} x_k e_k, \sum_{l=1}^{n} y_l e_l \right\rangle \\
&= \sum_{1 \leqslant k,l \leqslant n} \overline{x_k} y_l \langle e_k, e_l \rangle \\
&= \sum_{1 \leqslant k,l \leqslant n} \overline{x_k} \langle e_k, e_l \rangle y_l \\
&= {}^t\overline{\mathrm{Mat}_{\mathcal{B}}(v)} \left(\langle e_k, e_l \rangle\right)_{1 \leqslant k,l \leqslant n} \mathrm{Mat}_{\mathcal{B}}(w),
\end{aligned}$$

et on note $\mathrm{Mat}^{\mathrm{psh}}_{\mathcal{B}}(\langle \cdot,\cdot \rangle) := \left(\langle e_k, e_l \rangle\right)_{1 \leqslant k,l \leqslant n}$ la matrice représentative du produit scalaire hermitien $\langle \cdot,\cdot \rangle$ dans la base $\mathcal{B}$.

Exemple 3.4.1. On reprend les notations de l'exemple 3.3.8 : on note $\mathcal{B}$ la base $\{(1,1,0,0),(1,0,i,0),(0,0,1,1)\}$ du sous-espace vectoriel $F := \left\{(x,y,z,t) \in \mathbb{C}^4 \mid x - y + iz - it = 0\right\}$ de $\mathbb{C}^4$. Alors, en notant $\langle \cdot,\cdot \rangle$ la restriction du produit scalaire hermitien canonique sur $\mathbb{C}^4$ à F, on a

$$\mathrm{Mat}^{\mathrm{psh}}_{\mathcal{B}}(\langle \cdot,\cdot \rangle) = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & -i \\ 0 & i & 2 \end{pmatrix}.$$

À l'instar du cas euclidien, $\mathrm{Mat}^{\mathrm{psh}}_{\mathcal{B}}(\langle\cdot,\cdot\rangle) = I_n$ si et seulement si $\mathcal{B}$ est une base orthonormale de E.

Remarquons de plus que, comme la forme $\langle\cdot,\cdot\rangle$ est symétrique hermitienne, la matrice $\mathrm{Mat}^{\mathrm{psh}}_{\mathcal{B}}(\langle\cdot,\cdot\rangle)$ est hermitienne : une matrice A de $\mathrm{M}_n(\mathbb{C})$ est dite hermitienne si ${}^t\overline{A} = A$.

Par ailleurs, si $\mathcal{B}'$ est une autre de base de E :

Proposition 3.4.2. *On a*

$$\mathrm{Mat}^{\mathrm{psh}}_{\mathcal{B}'}(\langle\cdot,\cdot\rangle) = {}^t\overline{P_{\mathcal{B}\to\mathcal{B}'}}\,\mathrm{Mat}^{\mathrm{psh}}_{\mathcal{B}}(\langle\cdot,\cdot\rangle)\,P_{\mathcal{B}\to\mathcal{B}'}$$

Démonstration. Soient $v, w \in E$. Notons $X := \mathrm{Mat}_{\mathcal{B}}(v)$, $Y := \mathrm{Mat}_{\mathcal{B}}(w)$ et $A := \mathrm{Mat}^{\mathrm{psh}}_{\mathcal{B}}(\langle\cdot,\cdot\rangle)$, ainsi que $X' := \mathrm{Mat}_{\mathcal{B}'}(v)$, $Y' := \mathrm{Mat}_{\mathcal{B}'}(w)$ et $A' := \mathrm{Mat}^{\mathrm{psh}}_{\mathcal{B}'}(\langle\cdot,\cdot\rangle)$.

On a $X' = P_{\mathcal{B}'\to\mathcal{B}}X = {P_{\mathcal{B}\to\mathcal{B}'}}^{-1}X \Leftrightarrow X = P_{\mathcal{B}\to\mathcal{B}'}X'$ et $Y' = P_{\mathcal{B}'\to\mathcal{B}}Y = {P_{\mathcal{B}\to\mathcal{B}'}}^{-1}Y \Leftrightarrow Y = P_{\mathcal{B}\to\mathcal{B}'}Y'$. Ainsi,

$$\begin{array}{rcl}\langle v,w\rangle & = & {}^t\overline{X}AY \\ & = & {}^t\overline{(P_{\mathcal{B}\to\mathcal{B}'}X')}A\left(P_{\mathcal{B}\to\mathcal{B}'}Y'\right) \\ & = & {}^t\overline{X'}\left({}^t\overline{P_{\mathcal{B}\to\mathcal{B}'}}A\,P_{\mathcal{B}\to\mathcal{B}'}\right)Y'.\end{array}$$

En particulier, si l'on note $\mathcal{B}' = \{e'_1,\dots,e'_n\}$, on a, pour tous $k, l \in \{1,\dots,n\}$, $\langle e'_k, e'_l\rangle = {}^tX_k\left({}^t\overline{P_{\mathcal{B}\to\mathcal{B}'}}A\,P_{\mathcal{B}\to\mathcal{B}'}\right)X_l$ (avec les notations introduites dans la preuve de la proposition 2.5.4) et donc $A' = \left(\langle e'_k, e'_l\rangle\right)_{1\leqslant k,l\leqslant n} = {}^tP_{\mathcal{B}\to\mathcal{B}'}A\,\overline{P_{\mathcal{B}\to\mathcal{B}'}}$. □

3.5 Endomorphismes unitaires et matrices unitaires

Soit $(E, \langle\cdot,\cdot\rangle)$ un espace hermitien de dimension $n \in \mathbb{N}\backslash\{0\}$ et soit f un endomorphisme de E. De façon analogue au cadre euclidien, on peut considérer les endomorphismes qui préservent le produit scalaire hermitien :

Définition 3.5.1. *On dit que f est un endomorphisme unitaire de E si pour tous $v, w \in E$, $\langle f(v), f(w)\rangle = \langle v, w\rangle$.*

Cette dernière condition est équivalente à la préservation de la norme hermitienne $\|\cdot\|$ associée à $\langle\cdot,\cdot\rangle$:

Proposition 3.5.2. *f est un endomorphisme unitaire si et seulement si f conserve la norme hermitienne $\|\cdot\|$ si et seulement si f conserve la distance associée à la norme $\|\cdot\|$.*

Démonstration. On montre que si f conserve la norme hermitienne $\|\cdot\|$, alors f est unitaire. Le reste de la démonstration est tout à fait identique à la preuve de la proposition 2.6.3.

Supposons donc que f conserve $\|\cdot\|$ et soient $v, w \in E$, alors

$$\begin{array}{rcl}\langle f(v), f(w)\rangle & = & \dfrac{1}{2}\left(\|f(v)+f(w)\|^2 + i\|if(v)+f(w)\|^2 - (1+i)\|f(v)\|^2 - (1+i)\|f(w)\|^2\right)\ \text{(lemme 3.2.8)} \\ & = & \dfrac{1}{2}\left(\|f(v+w)\|^2 + i\|f(iv+w)\|^2 - (1+i)\|f(v)\|^2 - (1+i)\|f(w)\|^2\right)\ (f \text{ est } \mathbb{C}\text{-linéaire}) \\ & = & \dfrac{1}{2}\left(\|v+w\|^2 + i\|iv+w\|^2 - (1+i)\|v\|^2 - (1+i)\|w\|^2\right)\ (f \text{ préserve la norme hermitienne}) \\ & = & \langle v, w\rangle\ \text{(lemme 3.2.8).}\end{array}$$

□

À l'instar du cadre euclidien,

- l'endomorphisme f est unitaire si et seulement si f associe à toute base orthonormale de E une base orthonormale de E (la preuve de ce fait est tout à fait analogue à la preuve de la proposition 2.6.5),
- si f est unitaire, alors f est bijectif et f^{-1} est également unitaire (la démonstration est identique à celle du corollaire 2.6.7),
- la composition de deux endomorphismes unitaires est unitaire (la preuve est la même que celle de la proposition 2.6.16).

En particulier, l'ensemble des endomorphismes unitaires de E muni de la composition forme un groupe :

Définition 3.5.3. *L'ensemble, noté $\mathcal{U}\left(E, \langle\cdot,\cdot\rangle\right)$ ou simplement $\mathcal{U}(E)$ lorsque le contexte est clair, des endomorphismes unitaires de l'espace hermitien $(E, \langle\cdot,\cdot\rangle)$ est un sous-groupe du groupe linéaire $(\mathcal{GL}(E), \circ)$ des automorphismes linéaires de E. On appelle $(\mathcal{U}(E), \circ)$ le groupe unitaire de $(E, \langle\cdot,\cdot\rangle)$.*

On en vient maintenant à la caractérisation matricielle, dans une base orthonormale, du caractère unitaire d'un endomorphisme de E. Soit $\mathcal{B}$ une base orthonormale de E.

Proposition 3.5.4. *Notons $A := \mathrm{Mat}_{\mathcal{B}}(f)$ la matrice représentative de f dans la base $\mathcal{B}$. L'endomorphisme f de E est unitaire si et seulement si ${}^t\overline{A}A = I_n$.*

Démonstration. Notons $\mathcal{B} = \{e_1, \ldots, e_n\}$. L'endomorphisme f est alors unitaire si et seulement si, pour tous $k, l \in \{1, \ldots, n\}$, $\langle f(e_k), f(e_l)\rangle = \langle e_k, e_l\rangle = \delta_{k\,l}$. Or, pour tous $k, l \in \{1, \ldots, n\}$,

$$\langle f(e_k), f(e_l)\rangle = {}^t\overline{(AX_k)}(AX_l) = {}^t\overline{X_k}\left({}^t\overline{A}A\right)X_l$$

et donc f est unitaire si et seulement si ${}^t\overline{A}A = I_n$. □

On dira qu'une matrice A de $\mathrm{M}_n(\mathbb{C})$ est unitaire si elle vérifie ${}^t\overline{A}A = I_n$. Remarquons que

- une matrice de $\mathrm{M}_n(\mathbb{C})$ est unitaire si et seulement si les vecteurs colonnes qui la composent forment une base orthonormale de $\mathbb{C}^n$ muni du produit scalaire hermitien canonique,
- une matrice unitaire A est inversible et son inverse, également unitaire, est sa "transconjuguée" ${}^t\overline{A}$,
- si f est unitaire et si A désigne la matrice représentative de f dans la base orthonormale $\mathcal{B}$ de E alors $\mathrm{Mat}_{\mathcal{B}}\left(f^{-1}\right) = {}^t\overline{A}$ (on a $\mathrm{Mat}_{\mathcal{B}}\left(f^{-1}\right) = \mathrm{Mat}_{\mathcal{B}}\left(f\right)^{-1} = A^{-1} = {}^t\overline{A}$, car la matrice A est unitaire),
- le produit de matrices unitaires est unitaire.

On note $\mathrm{U}_n(\mathbb{C})$ le sous-groupe de $(\mathrm{GL}_n(\mathbb{C}), \cdot)$ formé par les matrices unitaires de $\mathrm{M}_n(\mathbb{C})$: on appelle $(\mathrm{U}_n(\mathbb{C}), \cdot)$ le groupe unitaire de $\mathrm{M}_n(\mathbb{C})$.

Exemple 3.5.5.
- La matrice
$$\frac{1}{2}\begin{pmatrix} 1 & -i & -1+i \\ i & 1 & 1+i \\ 1+i & -1+i & 0 \end{pmatrix}$$
de $\mathrm{M}_3(\mathbb{C})$ est unitaire.
- Toute matrice de passage d'une base orthonormale de $(E, \langle\cdot,\cdot\rangle)$ à une base orthonormale de $(E, \langle\cdot,\cdot\rangle)$ est unitaire.

Remarquons enfin que le déterminant d'une matrice unitaire, et donc d'un endomorphisme unitaire, est un nombre complexe de module 1 :

Proposition 3.5.6. *Soit $A \in \mathrm{U}_n(\mathbb{C})$, alors $|\det(A)| = 1$. Ainsi, si $f \in \mathcal{U}(E)$, $|\det(f)| = 1$.*

Démonstration. On a ${}^t\overline{A}A = I_n$ donc
$$1 = \det(I_n) = \det\left({}^t\overline{A}A\right) = \det\left({}^t\overline{A}\right)\det(A) = \det\left(\overline{A}\right)\det(A) = \overline{\det(A)}\det(A) = |\det(A)|.$$
□

Remarque 3.5.7. Le sous-ensemble de $\mathcal{U}(E)$ constitué des endomorphismes orthogonaux de déterminant 1 est un sous-groupe de $(\mathcal{U}(E), \circ)$ appelé groupe spécial unitaire de E et noté $\mathcal{SU}(E)$.

Le sous-ensemble de $\mathrm{U}_n(\mathbb{C})$ constitué des matrices unitaires de déterminant 1 est un sous-groupe de $(\mathrm{U}_n(\mathbb{C}), \cdot)$ appelé groupe spécial unitaire de $\mathrm{M}_n(\mathbb{C})$ et noté $\mathrm{SU}_n(\mathbb{C})$.

3.6 Décomposition QR d'une matrice à coefficients complexes inversible

Soit $n \in \mathbb{N}\backslash\{0\}$. Nous énonçons dans cette section l'analogue de la décomposition QR pour les matrices inversibles de $\mathrm{GL}_n(\mathbb{C})$: toute matrice inversible de $\mathrm{GL}_n(\mathbb{C})$ peut s'écrire comme le produit d'une matrice unitaire et d'une matrice triangulaire supérieure. Mais on peut être plus précis :

Théorème 3.6.1 (Décomposition QR d'une matrice inversible)**.** *Soit $n \in \mathbb{N}\backslash\{0\}$ et soit $A \in \mathrm{GL}_n(\mathbb{C})$. Il existe une unique matrice unitaire $Q \in \mathrm{U}_n(\mathbb{C})$ et une unique matrice $R \in \mathrm{GL}_n(\mathbb{C})$ triangulaire supérieure à coefficients diagonaux réels strictement positifs telles que $A = QR$. On appelle cette écriture la décomposition QR de A.*

Démonstration. L'existence d'une décomposition QR pour A est donnée, comme dans le cadre réel, par le procédé d'orthonormalisation de Gram-Schmidt (cf. théorème 3.3.7) : si $v_1, \ldots, v_n$ sont les vecteurs colonnes formant, dans l'ordre, les colonnes de la matrice A, on applique le procédé d'orthonormalisation de Gram-Schmidt à la base $\mathcal{B} := \{v_1, \ldots, v_n\}$ de $\mathbb{C}^n$ pour obtenir une base orthonormale $\mathcal{B}' = \{e_1, \ldots, e_n\}$ de $\mathbb{C}^n$ (par rapport au produit scalaire hermitien canonique).

De plus, par construction, pour tout $l \in \{1, \ldots, n\}$, $e_l = \sum_{k=1}^{l} \alpha_k v_k$ avec $\alpha_1, \ldots, \alpha_{l-1} \in \mathbb{C}$ et $\alpha_l \in]0; +\infty[$. Ainsi, la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$ de E est triangulaire supérieure et ses coefficients diagonaux sont réels strictement positifs. Il en est donc de même pour son inverse $R := P_{\mathcal{B}' \to \mathcal{B}}$ et, si l'on note Q la matrice de passage de la base canonique de $\mathbb{C}^n$ à la base $\mathcal{B}'$, Q est unitaire (les vecteurs colonnes qui composent Q forment une base orthonormale pour le produit scalaire hermitien canonique sur $\mathbb{C}^n$) et $A = QR$ (A peut être considérée comme la matrice de passage de la base canonique de $\mathbb{C}^n$ à $\mathcal{B}$).

La preuve de l'unicité de la décomposition QR de A est analogue à celle présentée dans la preuve du théorème 2.7.1. $\square$

Exemple 3.6.2. Considérons la matrice

$$A := \begin{pmatrix} 1 & 2 & 0 \\ i & 0 & 1 \\ 0 & -3i & 1 \end{pmatrix}$$

de $\mathrm{GL}_3(\mathbb{C})$ et notons $v_1 := (1, i, 0)$, $v_2 := (2, 0, -3i)$ et $v_3 := (0, 1, 1) \in \mathbb{C}^3$. On applique alors le procédé d'orthonormalisation de Gram-Schmidt à la base $\{v_1, v_2, v_3\}$ de $\mathbb{C}^3$. On pose

$$\begin{aligned}
\epsilon_1 &:= v_1 = (1, i, 0) \\
\epsilon_2 &:= v_2 - \frac{\overline{\langle v_2, \epsilon_1 \rangle}}{\|\epsilon_1\|^2}\epsilon_1 = v_2 - \frac{2}{2}\epsilon_1 = (1, -i, -3i) \\
\epsilon_3 &:= v_3 - \frac{\overline{\langle v_3, \epsilon_1 \rangle}}{\|\epsilon_1\|^2}\epsilon_1 - \frac{\overline{\langle v_3, \epsilon_2 \rangle}}{\|\epsilon_2\|^2}\epsilon_2 = v_3 + \frac{i}{2}\epsilon_1 - \frac{4i}{11}\epsilon_2 = \left(\frac{3i}{22}, \frac{3}{22}, -\frac{1}{11}\right) = \frac{1}{22}(3i, 3, -2) \\
\epsilon_3' &:= (3i, 3, -2) \\
e_1 &:= \frac{1}{\|\epsilon_1\|}\epsilon_1 = \frac{1}{\sqrt{2}}(1, i, 0) \\
e_2 &:= \frac{1}{\|\epsilon_2\|}\epsilon_2 = \frac{1}{\sqrt{11}}(1, -i, -3i) \\
e_3 &:= \frac{1}{\|\epsilon_3'\|}\epsilon_3' = \frac{1}{\sqrt{22}}(3i, 3, -2).
\end{aligned}$$

La matrice

$$Q := \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{11}} & \frac{3i}{\sqrt{22}} \\ \frac{i}{\sqrt{2}} & -\frac{i}{\sqrt{11}} & \frac{3}{\sqrt{22}} \\ 0 & -\frac{3i}{\sqrt{11}} & -\frac{2}{\sqrt{22}} \end{pmatrix}$$

est alors unitaire, et si l'on note

$$R := Q^{-1}A = {}^t\overline{Q}A = \begin{pmatrix} \sqrt{2} & \sqrt{2} & -\frac{i}{\sqrt{2}} \\ 0 & \sqrt{11} & \frac{4i}{\sqrt{11}} \\ 0 & 0 & \frac{1}{\sqrt{22}} \end{pmatrix},$$

l'égalité

$$A = QR = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{11}} & \frac{3i}{\sqrt{22}} \\ \frac{i}{\sqrt{2}} & -\frac{i}{\sqrt{11}} & \frac{3}{\sqrt{22}} \\ 0 & -\frac{3i}{\sqrt{11}} & -\frac{2}{\sqrt{22}} \end{pmatrix} \begin{pmatrix} \sqrt{2} & \sqrt{2} & -\frac{i}{\sqrt{2}} \\ 0 & \sqrt{11} & \frac{4i}{\sqrt{11}} \\ 0 & 0 & \frac{1}{\sqrt{22}} \end{pmatrix}$$

est la décomposition QR de A.

3.7 Endomorphismes hermitiens et matrices hermitiennes

Soit $(E, \langle \cdot, \cdot \rangle)$ un espace hermitien de dimension $n \in \mathbb{N} \backslash \{0\}$ et soit f un endomorphisme de E. L'analogue hermitien de la notion d'endomorphisme symétrique est la notion d'endomorphisme hermitien :

Définition 3.7.1. *On dit que l'endomorphisme f de E est hermitien si pour tous $v, w \in E$, $\langle f(v), w \rangle = \langle v, f(w) \rangle$.*

Si A est une matrice de $\mathrm{M}_n(\mathbb{C})$, on dit que A est hermitienne si A est égale à sa "transconjuguée" i.e. ${}^t\overline{A} = A$.

L'endomorphisme f est alors hermitien si et seulement sa matrice dans une base orthonormale de $(E, \langle \cdot, \cdot \rangle)$ est hermitienne :

Proposition 3.7.2. *Soit $\mathcal{B}$ une base orthonormale de E et notons $A := \mathrm{Mat}_{\mathcal{B}}(f)$ la matrice représentative de f dans $\mathcal{B}$. Alors f est hermitien si et seulement si la matrice A est hermitienne.*

Démonstration. Similairement au cadre euclidien (cf. preuve de la proposition 2.8.2), si $v, w \in E$ et si on note $X := \mathrm{Mat}_{\mathcal{B}}(v)$ et $Y := \mathrm{Mat}_{\mathcal{B}}(w)$, on a, comme $\mathcal{B}$ est une base orthonormale de E,

$$\langle f(v), w \rangle = {}^t\overline{(AX)}Y = {}^t\overline{X}{}^t\overline{A}Y \quad \text{et} \quad \langle v, f(w) \rangle = {}^t\overline{X}(AY) = {}^t\overline{X}AY.$$

Ainsi, pour tous $v, w \in E$, $\langle f(v), w \rangle = \langle v, f(w) \rangle$ si et seulement si ${}^t\overline{A} = A$ i.e. A est hermitienne. $\square$

Remarque 3.7.3. Les coefficients diagonaux d'une matrice hermitienne de $\mathrm{M}_n(\mathbb{C})$ sont nécessairement réels.

Nous allons à présent démontrer l'analogue hermitien du théorème spectral 2.8.4 :

Théorème 3.7.4 (Théorème spectral)**.** *On suppose que l'endomorphisme f est hermitien. Alors*

- *f est diagonalisable,*
- *les valeurs propres de f sont réelles,*
- *les sous-espaces propres de f sont orthogonaux deux à deux.*

En particulier, f est diagonalisable dans une base orthonormale de E.

Démonstration. On commence par montrer que les racines de χ_f (qui est un polynôme scindé car le corps de base est $\mathbb{C}$) sont toutes réelles. On reprend, en l'adaptant, l'argumentaire de la preuve du théorème 2.8.4 : soit $\mathcal{B}_0$ une base orthonormale de E, notons $A := \mathrm{Mat}_{\mathcal{B}_0}(f)$, soit $\lambda \in \mathrm{Sp}(A) = \mathrm{Sp}(f)$ et soit $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ un vecteur propre de A associé à λ. On applique alors la conjugaison complexe à l'égalité $AX = \lambda X$, de sorte que $\overline{A}\,\overline{X} = \overline{\lambda}\,\overline{X}$.

On considère maintenant l'égalité ${}^t(AX)\overline{X} = {}^tX\overline{A}\,\overline{X}$ (vérifiée car ${}^tA = \overline{A}$ car ${}^t\overline{A} = A$: A est la matrice représentative de l'endomorphisme hermitien f dans la base orthonormale $\mathcal{B}_0$) où l'on remplace AX par λX et $\overline{A}\,\overline{X}$ par $\overline{\lambda}\,\overline{X}$: on obtient ainsi

$$ {}^t(\lambda X)\overline{X} = {}^tX\overline{\lambda}\,\overline{X} \quad \text{i.e.} \quad \lambda \sum_{i=1}^{n} |x_i|^2 = \overline{\lambda} \sum_{i=1}^{n} |x_i|^2 \quad \text{i.e.} \quad \lambda = \overline{\lambda} \quad \text{i.e.} \quad \lambda \in \mathbb{R} $$

($\sum_{i=1}^{n} |x_i|^2 \neq 0$ car X, en tant que vecteur propre, n'est pas le vecteur nul).

Comme l'endomorphisme f est hermitien et ses valeurs propres sont réelles, la suite de la preuve est identique à celle du théorème 2.8.4. □

La version matricielle du théorème spectral 3.7.4 est le résultat suivant :

Corollaire 3.7.5. *Soit A une matrice hermitienne de $\mathrm{M}_n(\mathbb{C})$. Alors il existe une matrice unitaire $U \in \mathrm{U}_n(\mathbb{R})$ et une matrice diagonale $D \in \mathrm{M}_n(\mathbb{R})$ à coefficients réels telles que*

$$ D = U^{-1}AU = {}^t\overline{U}AU. $$

Démonstration. La preuve est similaire à la preuve du corollaire 2.8.7 : on utilise le fait qu'une matrice hermitienne de $\mathrm{M}_n(\mathbb{C})$ est la matrice représentative dans la base canonique de $\mathbb{C}^n$ d'un endomorphisme hermitien par rapport au produit scalaire hermitien canonique, et que la matrice de passage d'une base orthonormale à une base orthornomale est unitaire. □

Exemple 3.7.6. On considère la matrice $A := \begin{pmatrix} 0 & 1 & -i \\ 1 & 0 & -i \\ i & i & 2 \end{pmatrix} \in \mathrm{M}_3(\mathbb{C})$: A est hermitienne et on a

$$\begin{array}{rcl}
\chi_A & = & \begin{vmatrix} -X & 1 & -i \\ 1 & -X & -i \\ i & i & 2-X \end{vmatrix} \\
& \underset{C_1 \leftarrow C_1 - C_2}{=} & \begin{vmatrix} -X-1 & 1 & -i \\ 1+X & -X & -i \\ 0 & i & 2-X \end{vmatrix} \\
& = & (-1-X) \begin{vmatrix} 1 & 1 & -i \\ -1 & -X & -i \\ 0 & i & 2-X \end{vmatrix} \\
& \underset{L_1 \leftarrow L_2 + L_1}{=} & (-1-X) \begin{vmatrix} 1 & 1 & -i \\ 0 & 1-X & -2i \\ 0 & i & 2-X \end{vmatrix} \\
& = & (-1-X) \begin{vmatrix} 1-X & -2i \\ i & 2-X \end{vmatrix} \\
& = & (-1-X)[(1-X)(2-X)-2] \\
& = & (-1-X)\left(X^2 - 3X\right) \\
& = & (-1-X)(-X)(3-X).
\end{array}$$

Ainsi, $\mathrm{Sp}(A) = \{-1; 0; 3\}$. De plus, $E_{-1} = \mathrm{Vect}\{(1,-1,0)\} = \mathrm{Vect}\left\{\frac{1}{\sqrt{2}}(1,-1,0)\right\}$, $E_0 = \mathrm{Vect}\{(i,i,1)\} = \mathrm{Vect}\left\{\frac{1}{\sqrt{3}}(i,i,1)\right\}$ et $E_3 = \mathrm{Vect}\{(1,1,2i)\} = \mathrm{Vect}\left\{\frac{1}{\sqrt{6}}(1,1,2i)\right\}$.

Si l'on note alors

$$U := \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{2}} & \frac{i}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{3}} & \frac{2i}{\sqrt{6}} \end{pmatrix} \in \mathrm{M}_3(\mathbb{C}),$$

la matrice U est unitaire et on a

$$U^{-1}AU = {}^t\overline{U}AU = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Remarque 3.7.7. La réciproque du théorème 3.7.4 est également vraie, en ce sens que si l'endomorphisme f est diagonalisable dans une base orthonormale et si $\mathrm{Sp}(f) \subset \mathbb{R}$, alors f est hermitien. En effet, supposons qu'il existe une base orthonormale $\mathcal{B}$ de E et une matrice diagonale D de $\mathrm{M}_n(\mathbb{R}) \subset \mathrm{M}_n(\mathbb{C})$ telle que $\mathrm{Mat}_{\mathcal{B}}(f) = D$, alors

$${}^tD = D = \overline{D},$$

donc D est hermitienne, et en conséquence f est hermitien par la proposition 3.7.2 ($\mathcal{B}$ est une base orthonormale de E).

3.8 Diagonalisabilité des endomorphismes et matrices unitaires

Soit $(E, \langle \cdot, \cdot \rangle)$ un espace hermitien de dimension $n \in \mathbb{N} \backslash \{0\}$. Nous allons montrer dans cette section que tout endomorphisme unitaire de E est diagonalisable dans une base orthonormale et que ses valeurs propres sont toutes de module 1.

Précisément, soit f un endomorphisme de E, on a le résultat suivant :

Théorème 3.8.1. *Si f est unitaire alors*

- *f est diagonalisable,*
- *les valeurs propres de f sont de module 1,*
- *les sous-espaces propres de f sont orthogonaux deux à deux.*

En particulier, f est diagonalisable dans une base orthonormale de E.

Démonstration. On adapte le début de la preuve du théorème 2.9.1 du cadre euclidien.

Nous allons montrer par récurrence que pour tout $n \in \mathbb{N} \backslash \{0\}$, tout espace hermitien $(E, \langle \cdot, \cdot \rangle)$ de dimension n et tout endomorphisme unitaire f de E, les valeurs propres de f sont de module 1 et f est diagonalisable dans une base orthonormale de E (en particulier, les sous-espaces propres de f sont orthogonaux 2 à 2).

Pour $n = 1$, soit $E = \mathrm{Vect}\{v\}$ un $\mathbb{R}$-espace vectoriel de dimension 1 muni d'un produit scalaire hermitien $\langle \cdot, \cdot \rangle$ et soit f un endomorphisme unitaire de E. Il existe $\alpha \in \mathbb{C}$ tel que $f(v) = \alpha v$ et alors $\|f(v)\| = \|\alpha v\| = |\alpha| \|v\|$. De plus, comme f est unitaire, on a $\|f(v)\| = \|v\|$ donc $|\alpha| \|v\| = \|v\|$ et donc $|\alpha| = 1$ ($\|v\| \neq 0$ car $v \neq 0_E$ car v engendre E). Enfin la matrice de f dans la base $\{v\}$, ainsi que dans la base orthonormale $\left\{ \frac{v}{\|v\|} \right\}$, est (α).

Supposons maintenant la propriété vérifiée pour tout entier naturel non nul strictement plus petit que n avec $n \in \mathbb{N} \backslash \{0; 1\}$ fixé et considérons l'endomorphisme unitaire f de l'espace hermitien E de dimension n.

Soit $\lambda \in \mathrm{Sp}(f)$ et soit v un vecteur propre de f pour la valeur propre λ. On a d'une part $\|f(v)\| = \|\lambda v\| = |\lambda| \|v\|$ et, d'autre part, $\|f(v)\| = \|v\|$ car f est unitaire. Ainsi $|\lambda| \|v\| = \|v\|$ et donc $|\lambda| = 1$ (car v est un vecteur propre de f donc $v \neq 0_E$ donc $\|v\| \neq 0$). De plus, comme $F := \mathrm{Vect}\{v\}$ est stable par f (car v est un vecteur propre de f), l'orthogonal $F^\perp$ de F est également stable par f par la version hermitienne du lemme 2.9.3 (la preuve est identique) : comme $\dim\left(F^\perp\right) = n - 1 < n$, on peut alors appliquer l'hypothèse de récurrence à l'endomorphisme unitaire $f_{|F^\perp} : \begin{array}{ccc} F^\perp & \to & F^\perp \\ u & \mapsto & f(u) \end{array}$ de $F^\perp$ et obtenir l'existence d'une base orthonormale $\mathcal{B}_0$ de $F^\perp$ telle que

$$\mathrm{Mat}_{\mathcal{B}_0}\left(f_{|F^\perp}\right) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_{n-1} \end{pmatrix}$$

où les nombres complexes $\lambda_1, \dots, \lambda_{n-1}$ sont tous de module 1. Considérant l'égalité $F \oplus F^\perp = E$ (proposition 3.3.10) et la base orthonormale $\mathcal{B}' := \left\{ \frac{v}{\|v\|} \right\}$ de F, la famille $\mathcal{B} := \{B', \mathcal{B}_0\}$ est une base orthonormale de E et on a

$$\mathrm{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} \lambda & & & 0 \\ & \lambda_1 & & \\ & & \ddots & \\ 0 & & & \lambda_{n-1} \end{pmatrix}.$$

□

Remarque 3.8.2.

- La réciproque du théorème 3.8.1 est également vraie : si l'endomorphisme f est diagonalisable dans une base orthonormale et si ses valeurs propres sont toutes de module 1, alors f est unitaire.

  En effet, supposons qu'il existe une base orthonormale $\mathcal{B}$ de E et des nombres complexes $\lambda_1, \dots, \lambda_n$ de module 1 tels que

  $$\mathrm{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} =: D.$$

  Alors

  $${}^t\overline{D}D = \begin{pmatrix} \overline{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \overline{\lambda_n} \end{pmatrix} \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = \begin{pmatrix} |\lambda_1|^2 & & 0 \\ & \ddots & \\ 0 & & |\lambda_n|^2 \end{pmatrix} = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} = I_n.$$

  La matrice $D = \mathrm{Mat}_{\mathcal{B}}(f)$ est ainsi unitaire, et donc, par la proposition 3.5.4, l'endomorphisme f est unitaire (la base $\mathcal{B}$ de E est orthonormale).

- Si $A \in \mathrm{M}_n(\mathbb{C})$ est une matrice unitaire, alors il existe $U \in \mathrm{U}_n(\mathbb{C})$ et des nombres complexes $\lambda_1, \dots, \lambda_n$ tous de module 1 tels que

  $$U^{-1}AU = {}^t\overline{U}AU = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Exemple 3.8.3. On considère la matrice unitaire

$$A := \frac{1}{2} \begin{pmatrix} 1+i & 1-i & 0 \\ 1-i & 1+i & 0 \\ 0 & 0 & 2 \end{pmatrix} \in \mathrm{U}_3(\mathbb{C}).$$

Son polynôme caractéristique est

$$\begin{aligned}
\chi_A = \det(A - XI_3) &= \begin{vmatrix} \frac{1+i}{2} - X & \frac{1-i}{2} & 0 \\ \frac{1-i}{2} & \frac{1+i}{2} - X & 0 \\ 0 & 0 & 1-X \end{vmatrix} \\
&= (1-X)\left[\left(\frac{1+i}{2} - X\right)^2 - \left(\frac{1-i}{2}\right)^2\right] \\
&= (1-X)(i-X)(1-X) \\
&= (1-X)^2(i-X),
\end{aligned}$$

et on a

$$E_i = \operatorname{Ker} \begin{pmatrix} 1-i & 1-i & 0 \\ 1-i & 1-i & 0 \\ 0 & 0 & 2-2i \end{pmatrix} = \operatorname{Vect}\left\{\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}\right\} = \operatorname{Vect}\left\{\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}\right\}$$

et

$$E_1 = \operatorname{Ker} \begin{pmatrix} -1+i & 1-i & 0 \\ -1-i & 1+i & 0 \\ 0 & 0 & 0 \end{pmatrix} = \operatorname{Vect}\left\{\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right\} = \operatorname{Vect}\left\{\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right\}.$$

Ainsi, si l'on pose

$$U := \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

U est unitaire et on a

$$U^{-1}AU = {}^t\overline{U}AU = \begin{pmatrix} i & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Chapitre 4

Formes bilinéaires et formes quadratiques

4.1 Introduction

On étudie les propriétés algébriques et géométriques des formes bilinéaires, plus particulièrement des formes bilinéaires symétriques et des formes quadratiques associées. Plus générales que les produits scalaires euclidiens, les formes bilinéaires symétriques donnent lieu à une notion plus générale d'orthogonalité, non associée à une notion de distance. Néanmoins, une base orthogonale existe toujours dans le cadre de la dimension finie et on pourra déterminer une telle base orthogonale à l'aide de la méthode, dite de Gauss, de réduction d'une forme quadratique en somme de carrés de formes linéaires.

La fin de ce chapitre sera consacrée aux formes bilinéaires symétriques réelles pour lesquelles la relation d'ordre sur les nombres réels permet de définir un invariant (complet) : la signature.

4.2 Formes bilinéaires et formes bilinéaires symétriques

Soit $\mathbb{K}$ un corps commutatif de caractéristique différente de 2 (par exemple, $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$) et soit E un $\mathbb{K}$-espace vectoriel.

On considère les formes bilinéaires sur E i.e. les applications $\varphi : E \times E \to \mathbb{K}$ telles que, pour tous $v, w, v_1, v_2, w_1, w_2 \in E$ et tous $\lambda, \mu \in \mathbb{K}$,

$$\varphi(\lambda v_1 + \mu v_2, w) = \lambda\varphi(v_1, w) + \mu\varphi(v_2, w) \text{ et } \varphi(v, \lambda w_1 + \mu w_2) = \lambda\varphi(v, w_1) + \mu\varphi(v, w_2).$$

Soit $\varphi : E \times E \to \mathbb{K}$ une telle forme bilinéaire sur E.

Définition 4.2.1. *On dit que φ est <u>symétrique</u> si pour tous $v, w \in E$, $\varphi(w, v) = \varphi(v, w)$.*

Exemple 4.2.2. 1. Un produit scalaire sur un $\mathbb{R}$-espace vectoriel est, par définition, une forme bilinéaire symétrique sur ce dernier.

2. L'application
$$\phi : \begin{array}{ccc} \mathbb{R}^2 \times \mathbb{R}^2 & \to & \mathbb{R} \\ \big((x,y),(x',y')\big) & \mapsto & xy + x'y' \end{array}$$
n'est pas une forme bilinéaire (par exemple $\phi\big((0,0),(1,1)\big) = 1 \neq 0$).

3. L'application
$$\phi : \begin{array}{ccc} \mathbb{R}^2 \times \mathbb{R}^2 & \to & \mathbb{R} \\ \big((x,y),(x',y')\big) & \mapsto & xy' - x'y \end{array}$$
est une forme bilinéaire non symétrique : pour tous $(x,y),(x',y') \in \mathbb{R}^2$,
$$\phi\big((x',y'),(x,y)\big) = x'y - xy' = -\phi\big((x,y),(x',y')\big)$$
(ϕ est l'application déterminant dans la base canonique de $\mathbb{R}^2$).

4. Pour tout $n \in \mathbb{N}\backslash\{0\}$, l'application
$$\phi : \begin{array}{ccc} \mathbb{K}^n \times \mathbb{K}^n & \to & \mathbb{K} \\ \big((x_1,\dots,x_n),(y_1,\dots,y_n)\big) & \to & \displaystyle\sum_{k=1}^{n} x_k y_k \end{array}$$
est une forme bilinéaire symétrique sur $\mathbb{K}^n$.

5. Si $a, b \in \mathbb{R}$ et $\alpha \in \mathbb{R}$, l'application
$$\phi : \begin{array}{ccc} \mathcal{C}^0([a,b],\mathbb{R}) \times \mathcal{C}^0([a,b],\mathbb{R}) & \to & \mathbb{R} \\ (f,g) & \mapsto & \alpha \displaystyle\int_a^b f(t)g(t)\mathrm{d}t \end{array}$$
est une forme bilinéaire symétrique sur $\mathcal{C}^0([a,b],\mathbb{R})$.

6. Si $f_1,\dots,f_m,g_1,\dots,g_m \in \mathcal{L}(E,\mathbb{K})$, $\lambda_1,\dots,\lambda_m \in \mathbb{K}$, l'application
$$\phi : \begin{array}{ccc} E \times E & \to & \mathbb{K} \\ (v,w) & \mapsto & \displaystyle\sum_{k=1}^{m} \lambda_k f_k(v) g_k(w) \end{array}$$
est une forme bilinéaire sur E (non nécessairement symétrique).

On note $\mathcal{B}(E)$ l'ensemble des formes bilinéaires sur E et $\mathcal{S}(E)$ l'ensemble des formes bilinéaires symétriques sur E (on a $\mathcal{S}(E) \subset \mathcal{B}(E)$). Les ensembles $\mathcal{B}(E)$ et $\mathcal{S}(E)$ peuvent alors être chacun munis d'une structure de $\mathbb{K}$-espace vectoriel :

Proposition 4.2.3. *Les ensembles $\mathcal{B}(E)$ et $\mathcal{S}(E)$ sont des sous-espaces vectoriels de l'espace vectoriel des applications de $E \times E$ dans $\mathbb{K}$.*

Démonstration. L'application nulle $E \times E \to \mathbb{K}$; $(v, w) \mapsto 0$ est bilinéaire symétrique et est donc un élément de $\mathcal{S}(E) \subset \mathcal{B}(E)$.

Soient maintenant deux formes bilinéaires ϕ et ψ sur E et soient $\lambda, \mu \in \mathbb{K}$. Alors l'application $\lambda\phi + \mu\psi : E \times E \to \mathbb{K}$ est une forme bilinéaire sur E : pour tous $v, w, v_1, v_2, w_1, w_2 \in E$ et tous $\rho, \nu \in \mathbb{K}$,

$$\begin{array}{rcl}\big(\lambda\phi + \mu\psi\big)\big(\rho v_1 + \nu v_2, w\big) & = & \lambda\phi\big(\rho v_1 + \nu v_2, w\big) + \mu\psi\big(\rho v_1 + \nu v_2, w\big) \\ & = & \lambda\left[\rho\phi(v_1, w) + \nu\phi(v_2, w)\right] + \mu\left[\rho\psi(v_1, w) + \nu\psi(v_2, w)\right] \\ & = & \rho\left[\lambda\phi(v_1, w) + \mu\psi(v_1, w)\right] + \nu\left[\lambda\phi(v_2, w) + \mu\psi(v_2, w)\right] \\ & = & \rho\big(\lambda\phi + \mu\psi\big)(v_1, w) + \nu\big(\lambda\phi + \mu\psi\big)(v_2, w)\end{array}$$

et, de façon analogue,

$$\big(\lambda\phi + \mu\psi\big)\big(v, \rho w_1 + \nu w_2\big) = \rho\big(\lambda\phi + \mu\psi\big)(v, w_1) + \nu\big(\lambda\phi + \mu\psi\big)(v, w_2).$$

De plus, si ϕ et ψ sont symétriques, il en est de même pour $\lambda\phi + \mu\psi$: pour tous $v, w \in E$,

$$\begin{array}{rcl}\big(\lambda\phi + \mu\psi\big)(w, v) & = & \lambda\phi(w, v) + \mu\psi(w, v) \\ & = & \lambda\phi(v, w) + \mu\psi(v, w) \\ & = & \big(\lambda\phi + \mu\psi\big)(v, w).\end{array}$$

Ainsi, $\mathcal{B}(E)$ et $\mathcal{S}(E)$ sont bien des sous-espaces vectoriels de l'espace vectoriel des applications de $E \times E$ dans $\mathbb{K}$. □

Remarque 4.2.4. $\mathcal{S}(E)$ est un sous-espace vectoriel de $\mathcal{B}(E)$.

On continue cette section avec une identité valable pour toutes les formes bilinéaires :

Lemme 4.2.5. *Pour tous $v, w \in E$, on a*

$$\varphi(v + w, v + w) = \varphi(v, v) + \varphi(v, w) + \varphi(w, v) + \varphi(w, w).$$

Ainsi, si la forme bilinéaire φ est symétrique, on a, pour tous $v, w \in E$,

$$\varphi(v + w, v + w) = \varphi(v, v) + 2\varphi(v, w) + \varphi(w, w).$$

Démonstration. Soient $v, w \in E$, on a, par bilinéarité de φ,

$$\begin{array}{rcl}\varphi(v + w, v + w) & = & \varphi(v, v + w) + \varphi(w, v + w) \text{ (par linéarité à gauche de } \varphi) \\ & = & \varphi(v, v) + \varphi(v, w) + \varphi(w, v) + \varphi(w, w) \text{ (par linéarité à droite de } \varphi).\end{array}$$

□

Si E est de dimension finie, on va pouvoir "représenter" la forme bilinéaire φ de façon matricielle. Précisément, supposons dans la fin de cette section que E est de dimension finie $n \in \mathbb{N}\backslash\{0\}$ et soit $\mathcal{B} = \{e_1, \ldots, e_n\}$ une base de E. Si v et w sont deux vecteurs de E de coordonnées respectives $X := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \mathrm{Mat}_{\mathcal{B}}(v)$ et $Y := \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \mathrm{Mat}_{\mathcal{B}}(w)$ dans la base $\mathcal{B}$, on a alors, en notant A la matrice $\big(\varphi(e_k, e_l)\big)_{1\leqslant k,l\leqslant n}$,

$$\begin{aligned} \varphi(v,w) &= \varphi\left(\sum_{k=1}^n x_k e_k, \sum_{l=1}^n y_l e_l\right) \\ &= \sum_{1\leqslant k,l\leqslant n} x_k y_l\, \varphi(e_k, e_l) \\ &= \sum_{1\leqslant k,l\leqslant n} x_k\, \varphi(e_k, e_l)\, y_l \\ &= \sum_{k=1}^n x_k \left(\sum_{l=1}^n \varphi(e_k, e_l)\, y_l\right) \\ &= \sum_{k=1}^n x_k (AY)_k \\ &= {}^tXAY \end{aligned}$$

Cette écriture motive la définition suivante :

Définition 4.2.6. *On appelle <u>matrice représentative de la forme bilinéaire φ dans la base $\mathcal{B}$</u> la matrice*

$$\mathrm{Mat}^{\mathrm{fb}}_{\mathcal{B}}(\varphi) := \big(\varphi(e_k, e_l)\big)_{1\leqslant k,l\leqslant n} = \begin{pmatrix} \varphi(e_1,e_1) & \cdots & \varphi(e_1,e_n) \\ \vdots & & \vdots \\ \varphi(e_n,e_1) & \cdots & \varphi(e_n,e_n) \end{pmatrix}$$

Exemple 4.2.7.
1. La matrice représentative dans la base canonique $\mathcal{B}_{\mathrm{can}}$ de $\mathbb{R}^2$ de la forme bilinéaire ϕ de l'exemple 4.2.2 3. est
$$\mathrm{Mat}^{\mathrm{fb}}_{\mathcal{B}_{\mathrm{can}}}(\phi) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$
2. Pour tout $n \in \mathbb{N}\backslash\{0\}$, la matrice représentative dans la base canonique $\mathcal{B}_{\mathrm{can}}$ de $\mathbb{K}^n$ de la forme bilinéaire ϕ de l'exemple 4.2.2 4. est la matrice identité I_n de taille n.

Remarque 4.2.8. Attention à ne pas faire de confusions entre la matrice représentative d'une forme bilinéaire et la matrice représentative d'une application linéaire.

Soit $\mathcal{B}'$ une autre base de E, on a alors la relation de changement de base suivante pour la matrice représentative de la forme bilinéaire φ :

Proposition 4.2.9. *On a*

$$\mathrm{Mat}^{\mathrm{fb}}_{\mathcal{B}'}(\varphi) = {}^tP_{\mathcal{B}\to\mathcal{B}'}\, \mathrm{Mat}^{\mathrm{fb}}_{\mathcal{B}}(\varphi) P_{\mathcal{B}\to\mathcal{B}'}$$

Démonstration. La preuve est tout à fait identique à la preuve de la proposition 2.5.4 pour un produit scalaire. □

La symétrie d'une forme bilinéaire sur E est caractérisée par la symétrie de sa matrice représentative dans n'importe quelle base de E :

Proposition 4.2.10. *La forme bilinéaire φ est symétrique si et seulement si la matrice $A = \mathrm{Mat}_{\mathcal{B}}^{\mathrm{fb}}(\varphi)$ est symétrique i.e. ${}^tA = A$.*

Démonstration. Montrons que φ est symétrique si et seulement si pour tous $k, l \in \{1, \ldots, n\}$, $\varphi(e_k, e_l) = \varphi(e_l, e_k)$ ce qui montrera la proposition.

Si φ est symétrique i.e. pour tous $v, w \in E$, $\varphi(v, w) = \varphi(w, v)$, alors, en particulier, pour tous $k, l \in \{1, \ldots, n\}$, $\varphi(e_k, e_l) = \varphi(e_l, e_k)$.

Réciproquement, supposons que pour tous $k, l \in \{1, \ldots, n\}$, $\varphi(e_k, e_l) = \varphi(e_l, e_k)$, et soient $v = \sum_{k=1}^{n} x_k e_k$, $w = \sum_{l=1}^{n} y_l e_l$ deux vecteurs de E. Alors

$$\begin{aligned} \varphi(v, w) &= \varphi\left(\sum_{k=1}^{n} x_k e_k, \sum_{l=1}^{n} y_l e_l\right) \\ &= \sum_{1 \leqslant k, l \leqslant n} x_k y_l \varphi(e_k, e_l) \\ &= \sum_{1 \leqslant k, l \leqslant n} x_k y_l \varphi(e_l, e_k) \\ &= \varphi\left(\sum_{l=1}^{n} y_l e_l, \sum_{k=1}^{n} x_k e_k\right) \\ &= \varphi(w, v). \end{aligned}$$

Ainsi, φ est symétrique. □

Notons $\mathrm{S}_n(\mathbb{K}) := \left\{M \in \mathrm{M}_n(\mathbb{K}) \mid {}^tM = M\right\}$ le sous-espace vectoriel de $\mathrm{M}_n(\mathbb{K})$ des matrices symétriques ($I_n \in \mathrm{S}_n(\mathbb{K})$ et, si $M_1, M_2 \in \mathrm{S}_n(\mathbb{K})$ et $\lambda, \mu \in \mathbb{K}$, ${}^t(\lambda M_1 + \mu M_2) = \lambda {}^tM_1 + \mu {}^tM_2 = \lambda M_1 + \mu M_2$). La proposition précédente nous donne un isomorphisme linéaire entre $\mathcal{S}(E)$ et $\mathrm{S}_n(\mathbb{K})$:

Corollaire 4.2.11. *L'application de $\mathcal{S}(E)$ dans $\mathrm{S}_n(\mathbb{K})$ qui à toute forme bilinéaire symétrique sur E associe sa matrice représentative dans la base $\mathcal{B}$ de E est un isomorphisme linéaire (non canonique).*

Démonstration. Notons f l'application

$$\begin{array}{ccc} \mathcal{S}(E) & \to & \mathrm{S}_n(\mathbb{K}) \\ \phi & \mapsto & \mathrm{Mat}_{\mathcal{B}}^{\mathrm{fb}}(\phi) \end{array}.$$

Alors

- f est linéaire : si $\phi_1, \phi_2 \in \mathcal{S}(E)$, $\lambda, \mu \in \mathbb{K}$, alors, pour tous $k, l \in \{1, \ldots, n\}$, $(\lambda\phi_1 + \mu\phi_2)(e_k, e_l) = \lambda\phi_1(e_k, e_l) + \mu\phi_2(e_k, e_l)$,
- f est injective : $\mathrm{Mat}_{\mathcal{B}}^{\mathrm{fb}}(\phi) = 0_n$ ssi pour tous $k, l \in \{1, \ldots, n\}$, $\phi(e_k, e_l) = 0$ ssi ϕ est l'application nulle de $E \times E$ dans $\mathbb{K}$,
- f est surjective : si $M \in \mathrm{S}_n(\mathbb{K})$, l'application ϕ définie par, pour tous vecteurs $v, w \in E$ de coordonnées respectives $X := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $Y := \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ dans la base $\mathcal{B}$, $\phi(v, w) := {}^tXAY$, est une forme bilinéaire symétrique sur E,

donc f est bien un isomorphisme linéaire. $\square$

Remarque 4.2.12. Une matrice M de $\mathrm{M}_n(\mathbb{K})$ est dans $\mathrm{S}_n(\mathbb{K})$ si et seulement si elle est de la forme

$$M = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{1,2} & a_{2,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ a_{1,n} & \cdots & \cdots & a_{n,n} \end{pmatrix}$$

avec $a_{k,l} \in \mathbb{K}$, $1 \leqslant k \leqslant l \leqslant n$. En particulier,

$$\dim\left(\mathrm{S}_n(\mathbb{K})\right) = n + (n-1) + \cdots + 1 = \sum_{k=1}^{n} k = \frac{n(n+1)}{2},$$

et donc, en vertu du résultat précédent, $\dim\left(\mathcal{S}(E)\right) = \frac{n(n+1)}{2}$.

4.3 Formes quadratiques

Soient $\mathbb{K}$ un corps commutatif de caractéristique différente de 2 et E un espace vectoriel sur $\mathbb{K}$.

Définition 4.3.1. *On dit qu'une application $q : E \to \mathbb{K}$ est une forme quadratique sur E si*

1. *pour tout $v \in E$, pour tout $\lambda \in \mathbb{K}$, $q(\lambda v) = \lambda^2 q(v)$,*
2. *l'application*
$$\varphi : \begin{array}{ccc} E \times E & \to & \mathbb{K} \\ (v, w) & \mapsto & \frac{1}{2}\big(q(v+w) - q(v) - q(w)\big) \end{array}$$
est une forme bilinéaire symétrique sur E,

et, dans ce cas, l'application φ ci-dessus est appelée forme bilinéaire symétrique associée à q ou encore forme polaire de q.

Remarque 4.3.2.
1. Si q est une application de E dans $\mathbb{K}$, l'application φ ci-dessus définie est toujours symétrique : si $v, w \in E$,

$$\varphi(v,w) = \frac{1}{2}\big(q(v+w) - q(v) - q(w)\big) = \frac{1}{2}\big(q(w+v) - q(w) - q(v)\big) = \varphi(w,v).$$

En particulier, de par cette symétrie, l'application φ est linéaire à gauche ssi elle est linéaire à droite ssi elle est bilinéaire.

2. Avec les mêmes notations, si l'on suppose maintenant que $\forall v \in E, \forall \lambda \in \mathbb{K}, q(\lambda v) = \lambda^2 q(v)$, alors, pour tout $v \in E$,

$$\varphi(v,v) = \frac{1}{2}\big(q(2v) - q(v) - q(v)\big) = \frac{1}{2}\big(4q(v) - 2q(v)\big) = q(v).$$

3. Si q est une forme quadratique sur E de forme polaire φ, on a, pour tous $v, w \in E$,

$$q(v+w) = \varphi(v+w, v+w) = \varphi(v,v) + 2\varphi(v,w) + \varphi(w,w) = q(v) + 2\varphi(v,w) + q(w)$$

et

$$q(v-w) = q\big(v + (-w)\big) = q(v) + 2\varphi(v,-w) + q(-w) = q(v) - 2\varphi(v,w) + q(w).$$

Remarquons que si φ est une forme bilinéaire symétrique sur E, l'application

$$q : \begin{array}{ccc} E & \to & \mathbb{K} \\ v & \mapsto & \varphi(v,v) \end{array}$$

est une forme quadratique dont la forme polaire est φ (on appelle q <u>la forme quadratique de φ</u>). En effet,

1. pour tout $v \in E$, pour tout $\lambda \in \mathbb{K}$,

$$q(\lambda v) = \varphi(\lambda v, \lambda v) = \lambda^2 \varphi(v,v) = \lambda^2 q(v)$$

(par la bilinéarité de φ),

2. pour tous $v, w \in E$,

$$\begin{aligned} \frac{1}{2}\big(q(v+w) - q(v) - q(w)\big) &= \frac{1}{2}\big(\varphi(v+w, v+w) - \varphi(v,v) - \varphi(w,w)\big) \\ &= \frac{1}{2}\big(\varphi(v,v) + 2\varphi(v,w) + \varphi(w,w) - \varphi(v,v) - \varphi(w,w)\big) \\ &= \varphi(v,w). \end{aligned}$$

Ainsi,

Exemple 4.3.3. 1. L'application

$$\begin{array}{ccc} \mathbb{K}^n & \to & \mathbb{K} \\ (x_1, \dots, x_n) & \mapsto & \displaystyle\sum_{k=1}^{n} x_k^2 \end{array}$$

est la forme quadratique associée à la forme bilinéaire symétrique de l'exemple 4.2.2 4. Plus généralement, toute fonction polynomiale homogène de degré 2

$$q: \begin{array}{ccc} \mathbb{K}^n & \to & \mathbb{K} \\ (x_1, \dots, x_n) & \mapsto & \displaystyle\sum_{1 \leqslant k,l \leqslant n} a_{k,l} x_k x_l, \end{array}$$

avec, pour tous $k, l \in \{1, \dots, n\}$, $a_{k,l} \in \mathbb{K}$, est une forme quadratique :

- pour tout $(x_1, \dots, x_n) \in \mathbb{K}^n$ et tout $\lambda \in \mathbb{K}$,

$$q\big(\lambda(x_1, \dots, x_n)\big) = q(\lambda x_1, \dots, \lambda x_n) = \sum_{1 \leqslant k,l \leqslant n} a_{k,l} (\lambda x_k)(\lambda x_l) = \lambda^2 q(x_1, \dots, x_n),$$

- pour tous $v = (x_1, \dots, x_n), w = (y_1, \dots, y_n) \in \mathbb{K}^n$,

$$\begin{aligned} q(v+w) &= \sum_{1 \leqslant k,l \leqslant n} a_{k,l} (x_k + y_k)(x_l + y_l) \\ &= \sum_{1 \leqslant k,l \leqslant n} a_{k,l} x_k x_l + \sum_{1 \leqslant k,l \leqslant n} a_{k,l} y_k y_l + \sum_{1 \leqslant k,l \leqslant n} a_{k,l} (x_k y_l + x_l y_k) \\ &= q(v) + q(w) + \sum_{1 \leqslant k,l \leqslant n} a_{k,l} (x_k y_l + x_l y_k), \end{aligned}$$

ainsi

$$\frac{1}{2}\big(q(v+w) - q(v) - q(w)\big) = \frac{1}{2} \sum_{1 \leqslant k,l \leqslant n} a_{k,l} (x_k y_l + x_l y_k)$$

et l'application

$$\varphi: \begin{array}{ccc} E \times E & \to & \mathbb{K} \\ \big((x_1, \dots, x_n), (y_1, \dots, y_n)\big) & \mapsto & \displaystyle\frac{1}{2} \sum_{1 \leqslant k,l \leqslant n} a_{k,l} (x_k y_l + x_l y_k) \end{array}$$

est une forme bilinéaire (symétrique).

2. Toute application de la forme

$$q: \begin{array}{ccc} E & \to & \mathbb{K} \\ v & \mapsto & \displaystyle\sum_{k=1}^{m} \alpha_k \big(l_k(v)\big)^2 \end{array}$$

où, pour tout $k \in \{1, \dots, m\}$, $l_k \in \mathcal{L}(E, \mathbb{K})$ et $\alpha_k \in \mathbb{K}$, est une forme quadratique sur E :

- pour tout $v \in E$ et tout $\lambda \in \mathbb{K}$,

$$q(\lambda v) = \sum_{k=1}^{m} \alpha_k \big(l_k(\lambda v)\big)^2 = \sum_{k=1}^{m} \alpha_k \big(\lambda l_k(v)\big)^2 = \lambda^2 \sum_{k=1}^{m} \alpha_k \big(l_k(v)\big)^2 = \lambda^2 q(v),$$

- pour tous $v, w \in E$,

$$\begin{array}{rcl} q(v+w) & = & \displaystyle\sum_{k=1}^{m} \alpha_k \big(l_k(v+w)\big)^2 \\ & = & \displaystyle\sum_{k=1}^{m} \alpha_k \big(l_k(v) + l_k(w)\big)^2 \\ & = & \displaystyle\sum_{k=1}^{m} \alpha_k \left[\big(l_k(v)\big)^2 + 2l_k(v)l_k(w) + \big(l_k(w)\big)^2\right] \\ & = & q(v) + q(w) + \displaystyle\sum_{k=1}^{m} 2\alpha_k l_k(v) l_k(w) \end{array}$$

et l'application

$$\varphi : \begin{array}{ccc} E \times E & \to & \mathbb{K} \\ (v, w) & \mapsto & \displaystyle\sum_{k=1}^{m} \alpha_k l_k(v) l_k(w) \end{array}$$

est une forme bilinéaire symétrique sur E.

3. Si l_1 et l_2 sont deux formes linéaires de $\mathcal{L}(E, \mathbb{K})$, alors l'application

$$q : \begin{array}{ccc} E & \to & \mathbb{K} \\ w & \mapsto & l_1(w) l_2(w) \end{array}$$

est également une forme quadratique sur E :

- pour tout $v \in E$ et tout $\lambda \in \mathbb{K}$,

$$q(\lambda v) = l_1(\lambda v) l_2(\lambda v) = \lambda^2 l_1(v) l_2(v) = \lambda^2 q(v),$$

- pour tous $v, w \in E$,

$$\begin{array}{rcl} q(v+w) & = & l_1(v+w) l_2(v+w) \\ & = & l_1(v)l_2(v) + l_1(v)l_2(w) + l_1(w)l_2(v) + l_1(w)l_2(w) \\ & = & q(v) + q(w) + l_1(v)l_2(w) + l_1(w)l_2(v) \end{array}$$

et l'application

$$\varphi : \begin{array}{ccc} E \times E & \to & \mathbb{K} \\ (v, w) & \mapsto & \dfrac{1}{2}\big(l_1(v)l_2(w) + l_1(w)l_2(v)\big) \end{array}$$

est une forme bilinéaire symétrique sur E.

4. L'application
$$\begin{array}{ccc} \mathcal{C}^0([a,b],\mathbb{R}) & \to & \mathbb{R} \\ f & \mapsto & \alpha \displaystyle\int_a^b f(t)^2 \mathrm{d}t \end{array}$$
est la forme quadratique associée à la forme bilinéaire symétrique de l'exemple 4.2.2 5.

Proposition 4.3.4. *Soit q une forme quadratique sur E, de forme polaire φ. Pour tous $v, w \in E$, on a*

- $q(v+w) + q(v-w) = 2\big(q(v) + q(w)\big)$,
- $\varphi(v,w) = \dfrac{1}{4}\big(q(v+w) - q(v-w)\big)$.

Démonstration. Soient $v, w \in E$, on a

$$\begin{array}{rcl} q(v+w) + q(v-w) & = & q(v) + 2\varphi(v,w) + q(w) + q(v) - 2\varphi(v,w) + q(w) \text{ (par la remarque 4.3.2)} \\ & = & 2\big(q(v) + q(w)\big) \end{array}$$

et

$$\begin{array}{rcl} q(v+w) - q(v-w) & = & q(v) + 2\varphi(v,w) + q(w) - \big(q(v) - 2\varphi(v,w) + q(w)\big) \\ & = & 4\varphi(v,w) \end{array}$$

□

On note $\mathcal{Q}(E)$ l'ensemble des formes quadratiques sur E. Alors :

Proposition 4.3.5. *L'ensemble $\mathcal{Q}(E)$ est un sous-espace vectoriel de l'espace vectoriel des applications de E dans $\mathbb{K}$.*

Démonstration. L'application nulle $E \to \mathbb{K}$; $v \mapsto 0$ est une forme quadratique sur E et, si q_1, q_2 sont deux formes quadratiques sur E et $\mu, \nu \in \mathbb{K}$, alors $\mu q_1 + \nu q_2$ est une forme quadratique sur E :

1. pour tout $v \in E$, pour tout $\lambda \in \mathbb{K}$,
$$\begin{array}{rcl} (\mu q_1 + \nu q_2)(\lambda v) & = & \mu q_1(\lambda v) + \nu q_2(\lambda v) \\ & = & \mu\lambda^2 q_1(v) + \nu\lambda^2 q_2(v) \\ & = & \lambda^2 (\mu q_1 + \nu q_2)(v), \end{array}$$

2. si φ_1 et φ_2 désignent les formes polaires respectives de q_1 et q_2, alors, pour tous $v, w \in E$,
$$\begin{array}{rcl} \dfrac{1}{2}\big((\mu q_1 + \nu q_2)(v+w) - (\mu q_1 + \nu q_2)(v) - (\mu q_1 + \nu q_2)(w)\big) & = & \mu\dfrac{1}{2}\big(q_1(v+w) - q_1(v) - q_1(w)\big) \\ & & +\nu\dfrac{1}{2}\big(q_2(v+w) - q_2(v) - q_2(w)\big) \\ & = & \big(\mu\varphi_1 + \nu\varphi_2\big)(v,w) \end{array}$$
et l'application $\mu\varphi_1 + \nu\varphi_2$ est une forme bilinéaire symétrique sur E car $\mathcal{S}(E)$ est un espace vectoriel sur $\mathbb{K}$.

□

À toute forme quadratique sur E est associée, par définition, une forme bilinéaire symétrique sur E. Réciproquement, on a vu qu'à toute forme bilinéaire symétrique ϕ sur E pouvait être associée une forme quadratique sur E dont la forme polaire est ϕ.

Précisément :

Théorème 4.3.6. *L'application qui à toute forme quadratique sur E associe sa forme polaire est un isomorphisme linéaire (canonique) entre $\mathcal{Q}(E)$ et $\mathcal{S}(E)$.*

Démonstration. On note f l'application de $\mathcal{Q}(E)$ dans $\mathcal{S}(E)$ qui à une forme quadratique q associe sa forme polaire, et g l'application de $\mathcal{S}(E)$ dans $\mathcal{Q}(E)$ qui à une forme bilinéaire φ associe sa forme quadratique.

On a montré dans la preuve de la proposition 4.3.5 que si $q_1, q_2 \in \mathcal{Q}(E)$ et $\mu, \nu \in \mathbb{K}$, et si φ_1 et φ_2 désignaient les formes polaires respectives de q_1 et q_2, alors $\mu\varphi_1+\nu\varphi_2$ était la forme polaire de $\mu q_1+\nu q_2$, autrement dit $f\left(\mu q_1 + \nu q_2\right) = \mu f(q_1)+\nu f(q_2)$. Ainsi, f est une application linéaire.

On montre ensuite que f et g sont des applications réciproques l'une de l'autre i.e. $g \circ f = \mathrm{Id}_{\mathcal{Q}(E)}$ et $f \circ g = \mathrm{Id}_{\mathcal{S}(E)}$.

Soit tout d'abord $q \in \mathcal{Q}(E)$ et notons $\varphi := f(q)$ la forme polaire de q. Alors la forme quadratique $g(\varphi)$ de φ est q car il s'agit de l'application de E dans $\mathbb{K}$ qui à tout $v \in E$ associe $\varphi(v, v) = q(v)$ (cf. remarque 4.3.2 1.). Ainsi $g \circ f(q) = q$.

Soit ensuite $\varphi \in \mathcal{S}(E)$ et notons $q := g(\varphi)$ la forme quadratique de φ, alors la forme polaire $f(q)$ de q est φ car, pour tous $v, w \in E$, $\dfrac{1}{2}\big(q(v+w) - q(v) - q(w)\big) = \varphi(v, w)$. Ainsi $f \circ g(\varphi) = \varphi$.

En conséquence, $f : \mathcal{Q}(E) \to \mathcal{S}(E)$ est une application linéaire bijective et donc un isomorphisme linéaire. □

Remarque 4.3.7. Supposons que E soit de dimension finie $n \in \mathbb{N}\backslash\{0\}$.

1. On a alors $\dim\left(\mathcal{Q}(E)\right) = \dim\left(\mathcal{S}(E)\right) = \frac{n(n+1)}{2}$.

2. Soit $\mathcal{B} = \{e_1, \ldots, e_n\}$ une base de E et soit φ une forme bilinéaire symétrique sur E de forme polaire q. Notons, pour $k \in \{1, \ldots, n\}$, $a_k := \varphi(e_k, e_k)$ et, pour $k, l \in \{1, \ldots, n\}$ tels que $k < l$, $a_{k,l} := \varphi(e_k, e_l) = \varphi(e_l, e_k)$. Si $v = \sum_{k=1}^{n} x_k e_k,\ w = \sum_{k=1}^{n} y_k e_k \in E$, on a

$$\begin{aligned}
\varphi(v, w) &= \varphi\left(\sum_{k=1}^{n} x_k e_k, \sum_{l=1}^{n} y_l e_l\right) \\
&= \sum_{1 \leqslant k,l \leqslant n} x_k y_l \varphi(e_k, e_l) \\
&= \sum_{k=1}^{n} a_k\, x_k y_k + \sum_{1 \leqslant k < l \leqslant n} a_{k,l}\left(x_k y_l + x_l y_k\right)
\end{aligned}$$

et
$$q(v) = \sum_{k=1}^{n} a_k\, x_k^2 + 2 \sum_{1\leqslant k<l\leqslant n} a_{k,l}\, x_k x_l.$$

4.4 Orthogonalité et isotropie

Soient $\mathbb{K}$ un corps commutatif de caractéristique différente de 2, E un $\mathbb{K}$-espace vectoriel et φ une forme bilinéaire symétrique sur E de forme quadratique q.

On définit une relation d'orthogonalité relativement à φ :

Définition 4.4.1. *Soient $v, w \in E$. Comme φ est symétrique, $\varphi(v,w) = 0$ ssi $\varphi(w,v) = 0$, et on dit que les vecteurs v et w sont orthogonaux relativement à φ (ou φ-orthogonaux) si $\varphi(v,w) = 0$.*

Exemple 4.4.2.

1. Si l'on considère la forme bilinéaire symétrique
$$\phi : \begin{array}{ccc} \mathbb{R}^2\times\mathbb{R}^2 & \to & \mathbb{R} \\ \big((x,y),(x',y')\big) & \mapsto & xx'-yy' \end{array}$$
sur $\mathbb{R}^2$, les vecteurs $(1,2)$ et $(2,1)$ sont orthogonaux relativement à ϕ.

2. Si l'on considère la forme bilinéaire symétrique
$$\phi : \begin{array}{ccc} \mathbb{C}^2\times\mathbb{C}^2 & \to & \mathbb{R} \\ \big((x,y),(x',y')\big) & \mapsto & xx'+yy' \end{array}$$
sur $\mathbb{C}^2$, le vecteur $(i,1)$ est orthogonal à lui-même.

Soit A un sous-ensemble de E, on note
$$A^{\perp_\varphi} := \{v\in E \mid \forall w\in A,\ \varphi(v,w)=0\}$$
et on appelle orthogonal de A relativement à φ cet ensemble.

Exemple 4.4.3. Si l'on considère la forme bilinéaire symétrique
$$\phi : \begin{array}{ccc} \mathbb{R}^3\times\mathbb{R}^3 & \to & \mathbb{R} \\ \big((x,y,z),(x',y',z')\big) & \mapsto & 2xx'-3yy'-xz'-x'z \end{array}$$
sur $\mathbb{R}^3$,
$$\{(1,0,-1),(2,-1,0)\}^{\perp_\phi} = \left\{(x,y,z)\in\mathbb{R}^3 \mid 3x-z = 4x+3y-2z = 0\right\}.$$

On a les propriétés suivantes :

Proposition 4.4.4.

1. *$A^{\perp_\varphi}$ est un sous-espace vectoriel de E,*
2. *Soit B un sous-ensemble de E tel que $A\subset B$, alors $B^{\perp_\varphi}\subset A^{\perp_\varphi}$,*
3. *$A^{\perp_\varphi} = (\mathrm{Vect}(A))^{\perp_\varphi}$.*

Démonstration. 1. Le vecteur nul 0_E appartient à $A^{\perp_\varphi}$ car, pour tout $w \in A$, $\varphi(0_E, w) = 0$, et, si $v_1, v_2 \in A^{\perp_\varphi}$ et $\lambda, \mu \in \mathbb{K}$, alors, pour tout $w \in A$,

$$\varphi(\lambda v_1 + \mu v_2, w) = \lambda\varphi(v_1, w) + \mu\varphi(v_2, w) = 0$$

(car $v_1, v_2 \in A^{\perp_\varphi}$).

2. Si $v \in B^{\perp_\varphi}$, on a, pour tout $w \in A \subset B$, $\varphi(v, w) = 0$, donc $v \in A^{\perp_\varphi}$.

3. Par la propriété précédente, on a $(\mathrm{Vect}(A))^{\perp_\varphi} \subset A^{\perp_\varphi}$ (car $A \subset \mathrm{Vect}(A)$). Si maintenant $v \in A^{\perp_\varphi}$ et si $w = \sum_{k=1}^{m} \lambda_k w_k$ est une combinaison linéaire finie de vecteurs inclus dans A, alors

$$\varphi(v, w) = \varphi\left(v, \sum_{k=1}^{m} \lambda_k w_k\right) = \sum_{k=1}^{m} \lambda_k \varphi(v, w_k) = 0$$

(car $v \in A^{\perp_\varphi}$ et, pour tout $k \in \{1, \ldots, m\}$, $w_k \in A$), et donc $v \in (\mathrm{Vect}(A))^{\perp_\varphi}$.

□

On définit également une notion dite d'"isotropie" par rapport à φ :

Définition 4.4.5. *Soit $v \in E$. On dit que v est isotrope par rapport à φ si $\varphi(v, v) = q(v) = 0$. L'ensemble des vecteurs isotropes de E par rapport à φ est appelé cône isotrope de φ (ou de q) et noté C_φ.*

Exemple 4.4.6. Le cône isotrope de la forme bilinéaire symétrique

$$\phi : \begin{array}{ccc} \mathbb{R}^3 \times \mathbb{R}^3 & \to & \mathbb{R} \\ \big((x, y, z), (x', y', z')\big) & \mapsto & xx' + yy' - zz' \end{array}$$

est le cône de révolution d'équation $x^2 + y^2 = z^2$ dans $\mathbb{R}^3$.

Remarque 4.4.7. En général, le cône isotrope d'une forme quadratique n'est pas un sous-espace vectoriel de E. En reprenant les notations de l'exemple ci-dessus, même si les vecteurs $(1, 0, 1)$ et $(0, 1, 1)$ appartiennent au cône C_φ de φ, leur somme $(1, 0, 1) + (0, 1, 1) = (1, 1, 2)$ n'appartient pas à C_φ.

On définit d'autre part une notion de "noyau" pour φ :

Définition 4.4.8. *On appelle noyau de φ et on note Ker_φ l'orthogonal*

$$E^{\perp_\varphi} = \{v \in E \mid \forall w \in E,\ \varphi(v, w) = 0\}$$

de E relativement à φ (en particulier, Ker_φ est un sous-espace vectoriel de E).

On dit que la forme bilinéaire symétrique φ est non dégénérée si $\mathrm{Ker}_\varphi = \{0_E\}$, dégénérée sinon.

Exemple 4.4.9. 1. Tout produit scalaire est une forme bilinéaire symétrique non dégénérée.

2. Le noyau de la forme bilinéaire symétrique

$$\phi : \begin{array}{ccc} \mathbb{R}^2 \times \mathbb{R}^2 & \to & \mathbb{R} \\ \big((x,y),(x',y')\big) & \mapsto & xx' + xy' + x'y + yy' \end{array}$$

sur $\mathbb{R}^2$ est $\mathrm{Vect}\{(1,-1)\}$. En effet, si $(x,y) \in \mathrm{Ker}_\phi$ alors, pour tout $(x',y') \in \mathbb{R}^2$,

$$0 = \phi\big((x,y),(x',y')\big) = xx' + xy' + x'y + yy' = (x+y)(x'+y')$$

donc, en particulier, pour $x' = 1$ et $y' = 0$, on obtient $x + y = 0$ i.e. $y = -x$ et donc $(x,y) \in \mathrm{Vect}\{(1,-1)\}$. Réciproquement, pour tout $(x',y') \in E$, $\phi\big((1,-1),(x',y')\big) = x' + y' - x' - y' = 0$ donc $(1,-1) \in \mathrm{Ker}_\phi$ et donc $\mathrm{Vect}\{(1,-1)\} \subset \mathrm{Ker}_\phi$.

La forme bilinéaire ϕ est ainsi dégénérée.

Remarque 4.4.10. Attention à ne pas faire de confusions entre les notions de noyau d'une forme bilinéaire symétrique et de noyau d'une application linéaire !

Les notions de forme bilinéaire symétrique non dégénérée et de forme bilinéaire symétrique définie sont également à ne pas confondre :

Définition 4.4.11. *On dit que la forme bilinéaire symétrique φ est* <u>*définie*</u> *si, pour tout $v \in E$, $\varphi(v,v) = 0$ si et seulement si $v = 0_E$.*

Proposition 4.4.12. *Si φ est définie, alors φ est non dégénérée.*

Démonstration. Supposons que φ est définie et soit $v \in \mathrm{Ker}_\varphi$. En particulier, $\varphi(v,v) = 0$ et donc $v = 0_E$ (car φ a été supposée définie). Ainsi $\mathrm{Ker}_\varphi = \{0_E\}$ i.e. φ est non dégénérée. ☐

Remarque 4.4.13. La réciproque de la proposition 4.4.12 est fausse : par exemple, la forme bilinéaire symétrique

$$\phi : \begin{array}{ccc} \mathbb{R}^2 \times \mathbb{R}^2 & \to & \mathbb{R} \\ \big((x,y),(x',y')\big) & \mapsto & xx' - yy' \end{array}$$

sur $\mathbb{R}^2$ est non dégénérée et non définie. En effet, si $(x,y) \in \mathrm{Ker}_\phi$, alors en particulier $\phi\big((x,y),(1,0)\big) = x = 0$ et $\phi\big((x,y),(0,-1)\big) = y = 0$ donc $\mathrm{Ker}_\phi = \{0_E\}$. D'autre part, $\phi\big((1,1),(1,1)\big) = 0$ donc ϕ est non définie.

Supposons maintenant que E est de dimension finie $n \in \mathbb{N}\backslash\{0\}$. On peut alors caractériser matriciellement le caractère non dégénéré ou défini d'une forme bilinéaire symétrique. Introduisons tout d'abord la notion de rang d'une forme bilinéaire symétrique :

Définition 4.4.14. *Le* <u>*rang de la forme bilinéaire symétrique φ*</u> *est la quantité*

$$\dim(E) - \dim\left(\mathrm{Ker}_\varphi\right) = n - \dim\left(\mathrm{Ker}_\varphi\right)$$

que l'on note rg_φ.

Exemple 4.4.15. 1. Si ϕ est un produit scalaire sur E, alors la forme bilinéaire symétrique ϕ est non dégénérée et donc $\mathrm{rg}_\phi = \dim(E) = n$.

2. Le rang de la forme bilinéaire symétrique ϕ de l'exemple 4.4.9 2. est

$$\dim\left(\mathbb{R}^2\right) - \dim\left(\mathrm{Ker}_\phi\right) = 2 - 1 = 1.$$

Soit $\mathcal{B} = \{e_1, \ldots, e_n\}$ une base de E. Le rang de la forme bilinéaire symétrique φ est le rang de sa matrice représentative dans la base $\mathcal{B}$:

Proposition 4.4.16. *Notons* $A := \mathrm{Mat}^{\mathrm{fb}}_{\mathcal{B}}(\varphi)$. *Alors*

- *pour tout* $v \in E$, $v \in \mathrm{Ker}_\varphi$ *si et seulement si* $\mathrm{Mat}_{\mathcal{B}}(v) \in \mathrm{Ker}\, A$,
- $\mathrm{rg}_\varphi = \mathrm{rg}(A)$.

En particulier, φ *est non dégénérée si et seulement si* A *est inversible.*

Démonstration.

- Soit $v \in E$ et notons $X := \mathrm{Mat}_{\mathcal{B}}(v)$. Si $X \in \mathrm{Ker}\, A$, alors, pour tout $w \in E$,

$$\varphi(v, w) = \varphi(w, v) = {}^t\mathrm{Mat}_{\mathcal{B}}(w) A X = {}^t\mathrm{Mat}_{\mathcal{B}}(w) \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = 0.$$

Réciproquement, si $v \in \mathrm{Ker}_\varphi$, alors, pour tout $k \in \{1, \ldots, n\}$,

$$(AX)_k = {}^t\mathrm{Mat}_{\mathcal{B}}(e_k) A X = \varphi(e_k, v) = 0$$

et donc $AX = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$ i.e. $X \in \mathrm{Ker}\, A$.

- L'équivalence précédente donne un isomorphisme linéaire entre les espaces vectoriels Ker_φ et $\mathrm{Ker}\, A$ (l'application qui à un vecteur de Ker_φ associe le vecteur colonne de ses coordonnées dans la base $\mathcal{B}$). En particulier $\dim\left(\mathrm{Ker}_\varphi\right) = \dim\left(\mathrm{Ker}\, A\right)$ et donc

$$\mathrm{rg}_\varphi = \dim(E) - \dim\left(\mathrm{Ker}_\varphi\right) = n - \dim\left(\mathrm{Ker}\, A\right) = \mathrm{rg}(A).$$

□

Remarque 4.4.17. Avec les notations ci-dessus, la forme bilinéaire symétrique φ est définie si et seulement si la matrice $A = \mathrm{Mat}^{\mathrm{fb}}_{\mathcal{B}}(\varphi)$ est définie i.e. pour tout $X \in \mathrm{M}_{n,1}(\mathbb{K})$, ${}^tXAX = 0$ ssi $X = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$.

4.5 Bases orthogonales

Soient $\mathbb{K}$ un corps commutatif de caractéristique différente de 2, E un $\mathbb{K}$-espace vectoriel de dimension finie $n \in \mathbb{N}\backslash\{0\}$ et φ une forme bilinéaire symétrique sur E.

Soit $\mathcal{B} = \{e_1, \ldots, e_n\}$ une base de E.

Définition 4.5.1. *On dit que $\mathcal{B}$ est une base orthogonale de E relativement à φ (on dira aussi base φ-orthogonale) si les vecteurs e_k, $k \in \{1, \ldots, n\}$, sont deux à deux φ-orthogonaux i.e. pour tous $k, l \in \{1, \ldots, n\}$ tels que $k \neq l$, $\varphi(e_k, e_l) = 0$.*

Exemple 4.5.2.

1. La base canonique $\{(1,0),(0,1)\}$ de $\mathbb{R}^2$ est une base orthogonale relativement à la forme bilinéaire symétrique
$$\phi : \begin{array}{ccc} \mathbb{R}^2 \times \mathbb{R}^2 & \to & \mathbb{R} \\ \big((x,y),(x',y')\big) & \mapsto & xx' - yy' \end{array}$$

2. La famille $\{(1,-1),(1,0)\}$ de $\mathbb{R}^2$ est une base orthogonale relativement à la forme bilinéaire symétrique
$$\phi : \begin{array}{ccc} \mathbb{R}^2 \times \mathbb{R}^2 & \to & \mathbb{R} \\ \big((x,y),(x',y')\big) & \mapsto & xx' + xy' + x'y + yy' \end{array}$$

Remarque 4.5.3.

- Une famille de vecteurs deux à deux φ-orthogonaux peut ne pas être libre : avec les notations de l'exemple 4.5.2 2. ci-dessus, $\phi\big((1,-1),(2,-2)\big) = 0$. Cependant, si la forme bilinéaire symétrique φ est définie, alors toute famille orthogonale par rapport à φ est libre (cf. remarque 2.3.6 3.).
- La notion de "base orthonormale" n'a pas de sens pour une forme bilinéaire symétrique générale : une forme bilinéaire symétrique non définie positive ne peut donner lieu à une norme.

On peut caractériser l'orthogonalité de la base $\mathcal{B}$ relativement à φ à l'aide de la matrice représentative de φ dans la base $\mathcal{B}$:

Proposition 4.5.4. *La base $\mathcal{B}$ est une base φ-orthogonale si et seulement si la matrice $\mathrm{Mat}^{\mathrm{fb}}_{\mathcal{B}}(\varphi)$ est diagonale.*

Démonstration. La matrice $\mathrm{Mat}^{\mathrm{fb}}_{\mathcal{B}}(\varphi)$ est diagonale ssi pour tous $k, l \in \{1, \ldots, n\}$ tels que $k \neq l$, $\varphi(e_k, e_l) = 0$ ssi la base $\mathcal{B}$ est orthogonale relativement à φ. □

Remarque 4.5.5. Notons q la forme quadratique de φ.

- Si $\mathcal{B}$ est une base orthogonale pour φ, si $\mathrm{Mat}^{\mathrm{fb}}_{\mathcal{B}}(\varphi) = \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix}$ avec $a_1, \ldots, a_n \in \mathbb{K}$, et si $v = \displaystyle\sum_{k=1}^{n} x_k e_k$, $w = \displaystyle\sum_{k=1}^{n} y_k e_k \in E$, on a
$$\varphi(v,w) = a_1 x_1 y_1 + \cdots + a_n x_n y_n \quad \text{et} \quad q(v) = a_1 x_1^2 + \cdots + a_n x_n^2.$$

- Réciproquement, supposons qu'il existe $a_1, \ldots, a_n \in \mathbb{K}$ tels que, pour tout $v = \sum_{k=1}^{n} x_k e_k \in E$, $q(v) = a_1 x_1^2 + \cdots + a_n x_n^2$, alors la base $\mathcal{B} = \{e_1, \ldots, e_n\}$ est orthogonale pour φ. En effet, dans ce cas, pour tous $k, l \in \{1, \ldots, n\}$ tels que $k \neq l$,

$$\varphi(e_k, e_l) = \frac{1}{2}\big(q(e_k + e_l) - q(e_k) - q(e_l)\big) = \frac{1}{2}\big(a_k + a_l - a_k - a_l\big) = 0.$$

Proposition 4.5.6. *On suppose que $\mathcal{B} = \{e_1, \ldots, e_n\}$ est une base φ-orthogonale. Alors la sous-famille $\{e_k \mid q(e_k) = 0\}$ est une base de Ker_φ et le cardinal de la sous-famille $\{e_k \mid q(e_k) = 0\}$ (i.e. le nombre de coefficients diagonaux non nuls dans la matrice diagonale $\mathrm{Mat}_{\mathcal{B}}^{\mathrm{fb}}(\varphi)$) est le rang rg_φ de φ.*

Démonstration. Soit $k \in \{1, \ldots, n\}$, alors, pour tout $l \in \{1, \ldots, n\}$ tel que $l \neq k$, $\varphi(e_k, e_l) = 0$ (car $\{e_1, \ldots, e_n\}$ est une base φ-orthogonale). Si de plus $q(e_k) = \varphi(e_k, e_k) = 0$ alors e_k est φ-orthogonal à tous les vecteurs de E (par bilinéarité de φ) donc $e_k \in E^{\perp_\varphi} = \mathrm{Ker}_\varphi$. Ainsi, la famille $\{e_k \mid q(e_k) = 0\}$ est une famille libre de Ker_φ.

Montrons maintenant qu'elle engendre Ker_φ : soit $v \in \mathrm{Ker}_\varphi$ et écrivons $v = \sum_{k=1}^{n} x_k e_k$ avec $x_1, \ldots, x_n \in \mathbb{K}$ ($\mathcal{B}$ est une base de E). Pour tout $l \in \{1, \ldots, n\}$, on a

$$\begin{aligned} 0 &= \varphi(v, e_l) \text{ (car } v \in \mathrm{Ker}_\varphi) \\ &= \varphi\left(\sum_{k=1}^{n} x_k e_k, e_l\right) \\ &= \sum_{k=1}^{n} x_k \varphi(e_k, e_l) \\ &= x_l q(e_l) \text{ (car } \{e_1, \ldots, e_n\} \text{ est une famille orthogonale).} \end{aligned}$$

Ainsi, pour tout $l \in \{1, \ldots, n\}$ tel que $q(e_l) \neq 0$, $x_l = 0$ et v est donc une combinaison linéaire des vecteurs e_k tels que $q(e_k) = 0$.

Au total, la famille des vecteurs e_k tels que $q(e_k) = 0$ est bien une base (φ-orthogonale) de Ker_φ. □

Un résultat important de la théorie des formes bilinéaires symétriques sur un espace vectoriel de dimension finie est qu'il existe toujours des bases orthogonales :

Théorème 4.5.7. *Il existe une base de E orthogonale relativement à φ.*

Démonstration. On montre le résultat par récurrence sur la dimension n de E : précisément, on montre que pour tout $n \in \mathbb{N}\backslash\{0\}$, pour tout espace vectoriel E de dimension finie n, pour toute forme bilinéaire symétrique φ sur E, il existe une base de E orthogonale relativement à φ.

Si $n = 1$, la matrice de toute forme bilinéaire (symétrique) sur un espace vectoriel de dimension 1, étant de taille 1, est diagonale.

Supposons maintenant la propriété vérifiée au rang $n-1$ pour $n \in \mathbb{N}\backslash\{0;1\}$ fixé et considérons notre forme bilinéaire symétrique φ sur notre $\mathbb{K}$-espace vectoriel E de dimension n. Si φ est la forme linéaire nulle, alors la matrice représentative de φ dans toute base est la matrice nulle de taille n et est en particulier diagonale. Supposons donc que φ n'est pas la forme bilinéaire nulle et soit $v \in E$ tel que $q(v) \neq 0$: par l'isomorphisme linéaire entre $\mathcal{Q}(E)$ et $\mathcal{S}(E)$ (cf. théorème 4.3.6), φ est la forme bilinéaire symétrique nulle ssi q est la forme quadratique nulle.

On considère alors l'orthogonal

$$F := \{v\}^{\perp_\varphi} = \{w \in E \mid \varphi(v,w) = 0\}$$

de $\{v\}$ par rapport à φ : il s'agit du noyau de l'application linéaire

$$\varphi(v,\cdot) : \begin{array}{ccc} E & \to & \mathbb{K} \\ w & \mapsto & \varphi(v,w) \end{array}.$$

Comme $\varphi(v,v) = q(v) \neq 0$, l'application linéaire $\varphi(v,\cdot)$ est surjective et donc

$$\dim(F) = \dim\big(\mathrm{Ker}\ \varphi(v,\cdot)\big) = \dim(E) - \dim \mathbb{K} = n-1.$$

De plus, les sous-espaces vectoriels $\mathrm{Vect}\{v\}$ et F de E sont en somme directe (car $v \notin \dim \mathrm{Ker}\ \varphi(v,\cdot)$) donc $E = \mathrm{Vect}\{v\} \oplus F$.

On applique ensuite l'hypothèse de récurrence à l'espace vectoriel F de dimension $n-1$ et à la forme bilinéaire symétrique restreinte $\varphi_{|F\times F} : F \times F \to \mathbb{K}$; $(v,w) \mapsto \varphi(v,w)$: il existe une base $\{v_2,\ldots,v_n\}$ de F orthogonale relativement à $\varphi_{|F\times F}$. Si l'on note enfin $v_1 := v$, la famille $\{v_1,v_2,\ldots,v_n\}$ est une base de E (car $E = \mathrm{Vect}\{v\} \oplus F$), orthogonale relativement à φ (car pour tous $k,l \in \{2,\ldots,n\}$ tels que $k \neq l$, $\varphi(v_k,v_l) = \varphi_{F\times F}(v_k,v_l) = 0$, et, pour tout $k \in \{2,\ldots,n\}$, $\varphi(v_1,v_k) = \varphi(v,v_k) = 0$). □

Remarque 4.5.8. 1. On ne peut pas toujours compléter une famille libre de vecteurs deux à deux φ-orthogonaux en une base φ-orthogonale. Par exemple, si $E = \mathbb{R}^2$ et $\varphi = \mathbb{R}^2 \to \mathbb{R}^2$; $\big((x,y),(x',y')\big) \mapsto \frac{1}{2}\big(xy' + x'y\big)$, alors

$$\{(1,0)\}^{\perp_\varphi} = \big\{(x,y) \in \mathbb{R}^2 \mid 1 \times y + x \times 0 = y = 0\big\} = \mathrm{Vect}\{(1,0)\} :$$

il est donc impossible de compléter la famille libre $\{(1,0)\}$ en une base orthogonale relativement à φ : tout vecteur φ-orthogonal à $(1,0)$ est dans la droite vectorielle engendrée par $(1,0)$.

2. L'existence d'une base orthogonale pour toute forme bilinéaire symétrique permet de montrer qu'une forme bilinéaire symétrique sur un $\mathbb{C}$-espace vectoriel de dimension finie au moins égale à 2 n'est jamais définie.

4.6 Réduction d'une forme quadratique en somme de carrés et méthode de Gauss

Dans cette section, soit $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ et considérons un $\mathbb{K}$-espace vectoriel E de dimension finie et une forme bilinéaire symétrique φ sur E de forme quadratique q. Nous allons décrire une démarche systématique pour déterminer une base φ-orthogonale.

Le point de départ de cette démarche est la réduction d'une forme quadratique sur $\mathbb{K}$ en "somme de carrés" dite méthode de Gauss.

Notons $n := \dim(E)$ et commençons par considérer une base $\{e_1, \ldots, e_n\}$ de E. Notons, pour $k \in \{1, \ldots, n\}$, $a_k := \varphi(e_k, e_k)$ et, pour $k, l \in \{1, \ldots, n\}$ tels que $k < l$, $a_{k,l} := \varphi(e_k, e_l) = \varphi(e_l, e_k)$. Si $v = \sum_{k=1}^{n} x_k e_k \in E$, on a alors

$$q(v) = \sum_{k=1}^{n} a_k\, x_k^2 + 2 \sum_{1 \leqslant k < l \leqslant n} a_{k,l}\, x_k x_l$$

(cf. remarque 4.3.7 2.).

Théorème 4.6.1. *Il existe n formes linéaires $l_1, \ldots, l_n \in \mathcal{L}(E, \mathbb{K})$ linéairement indépendantes (en tant que vecteurs de $\mathcal{L}(E, \mathbb{K})$) et $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ tels que, pour tout $v \in E$,*

$$q(v) = \sum_{k=1}^{n} \alpha_k \big(l_k(v)\big)^2.$$

Démonstration. On montre le résultat par récurrence sur n.

Si $n = 1$, toute base d'un espace vectoriel de dimension 1 est une base orthogonale relativement à toute forme bilinéaire (symétrique) sur cet espace et la forme quadratique correspondante s'exprime alors comme un scalaire fois le carré d'une forme linéaire.

Supposons maintenant la propriété vérifiée à tout rang $k \in \{1, \ldots, n-1\}$ pour $n \in \mathbb{N}\backslash\{0; 1\}$ fixé et considérons notre forme quadratique q sur notre $\mathbb{R}$-espace vectoriel E de dimension n. On suppose q non identiquement nulle (la forme quadratique nulle sur E peut s'écrire comme la combinaison linéaire nulle de n'importe quelle famille linéairement indépendante de formes linéaires sur E).

Reprenons, avec les notations ci-dessus, l'expression

$$q(v) = \sum_{k=1}^{n} a_k\, x_k^2 + 2 \sum_{1 \leqslant k < l \leqslant n} a_{k,l}\, x_k x_l$$

pour $v = \sum_{k=1}^{n} x_k e_k \in E$.

On différencie deux cas :

- S'il existe $k \in \{1, \ldots, n\}$ tel que $a_k = \varphi(e_k, e_k) \neq 0$: On commence par permuter le vecteur e_k et le vecteur e_1 de la base $\mathcal{B}$ de façon à ce que $a_1 \neq 0$, et on a alors

$$\begin{aligned}
q(v) &= a_1 x_1^2 + \sum_{k=2}^{n} a_k\, x_k^2 + 2 \sum_{1 \leqslant k < l \leqslant n} a_{k,l}\, x_k x_l \\
&= a_1 x_1^2 + 2x_1 \sum_{2 \leqslant l \leqslant n} a_{1,l}\, x_l + \sum_{k=2}^{n} a_k\, x_k^2 + 2 \sum_{2 \leqslant k < l \leqslant n} a_{k,l}\, x_k x_l \\
&= a_1 x_1^2 + 2x_1 f\left(\sum_{k=2}^{n} x_k e_k\right) + Q\left(\sum_{k=2}^{n} x_k e_k\right)
\end{aligned}$$

où, si l'on note $F := \mathrm{Vect}\{e_2, \ldots, e_n\}$, f est la forme linéaire

$$\begin{array}{ccc} F & \to & \mathbb{K} \\ \sum\limits_{k=2}^{n} y_k e_k & \mapsto & \sum\limits_{2 \leqslant k \leqslant n} a_{1,k}\, y_k \end{array}$$

sur F, et Q est la forme quadratique

$$\begin{array}{ccc} F & \to & \mathbb{K} \\ \sum\limits_{k=2}^{n} y_k e_k & \mapsto & \sum\limits_{k=2}^{n} a_k\, y_k^2 + 2 \sum\limits_{2 \leqslant k < l \leqslant n} a_{k,l}\, y_k y_l \end{array}$$

sur F. On écrit ensuite, en notant $w := \sum\limits_{k=2}^{n} x_k e_k$,

$$\begin{aligned} q(v) &= a_1 x_1^2 + 2 x_1 f(w) + Q(w) \\ &= a_1 \left[x_1 + \frac{f(w)}{a_1} \right]^2 - \frac{\big(f(w)\big)^2}{a_1} + Q(w). \end{aligned}$$

Or l'application

$$l_1 : \begin{array}{ccc} E & \to & \mathbb{K} \\ v = \sum\limits_{k=1}^{n} x_k e_k & \mapsto & x_1 + \dfrac{f\left(\sum\limits_{k=2}^{n} x_k e_k\right)}{a_1} \end{array}$$

est une forme linéaire sur E, et l'application

$$\begin{array}{ccc} F & \to & \mathbb{K} \\ w & \mapsto & -\frac{\big(f(w)\big)^2}{a_1} + Q(w) \end{array}$$

est une forme quadratique sur F : comme $\dim(F) = n-1$, par hypothèse de récurrence, il existe $n-1$ formes linéaires $\rho_2, \ldots, \rho_n \in \mathcal{L}(F, \mathbb{K})$ sur F et des scalaires $\alpha_2, \ldots, \alpha_n \in \mathbb{K}$ tels que, pour tout $w \in F$,

$$-\frac{\big(f(w)\big)^2}{a_1} + Q(w) = \sum_{k=2}^{n} \alpha_k \big(\rho_k(w)\big)^2.$$

Si, pour tout $k \in \{2, \ldots, n\}$, on note alors

$$l_k : \begin{array}{ccc} E & \to & \mathbb{K} \\ v = \sum\limits_{k=1}^{n} x_k e_k & \mapsto & \rho_k\left(\sum\limits_{k=2}^{n} x_k e_k\right) \end{array},$$

on obtient finalement l'égalité

$$q(v) = a_1\big(l_1(v)\big)^2 + \sum_{k=2}^{n} \alpha_k\big(l_k(w)\big)^2.$$

De plus, la famille $\{l_1, l_2, \ldots, l_n\}$ de $\mathcal{L}(E, \mathbb{K})$ est une famille libre car la famille $\{l_2, \ldots, l_n\}$ est libre et l'expression de la forme linéaire l_1 dépend de la coordonnée x_1 ce qui n'est pas le cas pour $l_2, \ldots, l_n$.

- Si pour tout $k \in \{1, \ldots, n\}$, $a_k = 0$: Alors on a

$$q(v) = 2 \sum_{1 \leqslant k < l \leqslant n} a_{k,l}\, x_k x_l.$$

Comme q n'est pas la forme quadratique nulle par hypothèse, il existe $k, l \in \{1, \ldots, n\}$ avec $k < l$ tels que $a_{k,l} \neq 0$. On permute alors les vecteurs de la base $\mathcal{B}$ de façon à ce que $a_{1,2} \neq 0$ et on écrit

$$\begin{aligned} q(v) &= 2a_{1,2}\, x_1 x_2 + 2x_1 \sum_{k=3}^{n} a_{1,k}\, x_k + 2x_2 \sum_{k=3}^{n} a_{2,k}\, x_k + 2 \sum_{3 \leqslant k < l \leqslant n} a_{k,l}\, x_k x_l \\ &= 2a_{1,2}\, x_1 x_2 + 2x_1 f_1\left(\sum_{k=3}^{n} x_k e_k\right) + 2x_2 f_2\left(\sum_{k=3}^{n} x_k e_k\right) + Q\left(\sum_{k=3}^{n} x_k e_k\right) \end{aligned}$$

où, si l'on note $G := \mathrm{Vect}\{e_3, \ldots, e_n\}$ et si $l \in \{1, 2\}$, f_l est la forme linéaire

$$\begin{array}{ccc} G & \to & \mathbb{K} \\ \displaystyle\sum_{k=3}^{n} y_k e_k & \mapsto & \displaystyle\sum_{k=3}^{n} a_{l,k}\, y_k \end{array}$$

sur G, et Q est la forme quadratique

$$\begin{array}{ccc} G & \to & \mathbb{K} \\ \displaystyle\sum_{k=3}^{n} y_k e_k & \mapsto & \displaystyle 2 \sum_{3 \leqslant k < l \leqslant n} a_{k,l}\, y_k y_l \end{array}$$

sur G. Notant $w := \displaystyle\sum_{k=3}^{n} x_k e_k$, on écrit ensuite

$$\begin{aligned} q(v) &= 2a_{1,2}\, x_1 x_2 + 2x_1 f_1(w) + 2x_2 f_2(w) + Q(w) \\ &= 2a_{1,2}\left[x_1 x_2 + \frac{x_1}{a_{1,2}} f_1(w) + \frac{x_2}{a_{1,2}} f_2(w)\right] + Q(w) \\ &= 2a_{1,2}\left[\left(x_1 + \frac{f_2(w)}{a_{1,2}}\right)\left(x_2 + \frac{f_1(w)}{a_{1,2}}\right) - \frac{f_1(w) f_2(w)}{a_{1,2}^2}\right] + Q(w) \\ &= 2a_{1,2}\left(x_1 + \frac{f_2(w)}{a_{1,2}}\right)\left(x_2 + \frac{f_1(w)}{a_{1,2}}\right) - 2\frac{f_1(w) f_2(w)}{a_{1,2}} + Q(w) \\ &= \frac{a_{1,2}}{2}\left[\left(x_1 + x_2 + \frac{f_1(w) + f_2(w)}{a_{1,2}}\right)^2 - \left(x_1 - x_2 + \frac{f_2(w) - f_1(w)}{a_{1,2}}\right)^2\right] - 2\frac{f_1(w) f_2(w)}{a_{1,2}} + Q(w) \end{aligned}$$

où l'on a utilisé le fait que, pour tous $a, b \in \mathbb{K}$, $ab = \frac{1}{4}\left((a+b)^2 - (a-b)^2\right)$. Or les applications

$$l_1 : \begin{array}{ccc} E & \to & \mathbb{K} \\ v = \sum_{k=1}^{n} x_k e_k & \mapsto & x_1 + x_2 + \dfrac{f_1(w) + f_2(w)}{a_{1,2}} \end{array}$$

et

$$l_2 : \begin{array}{ccc} E & \to & \mathbb{K} \\ v = \sum_{k=1}^{n} x_k e_k & \mapsto & x_1 - x_2 + \dfrac{f_2(w) - f_1(w)}{a_{1,2}} \end{array}$$

sont des formes linéaires sur E, et l'application

$$\begin{array}{ccc} G & \to & \mathbb{K} \\ w & \mapsto & -2\frac{f_1(w)f_2(w)}{a_{1,2}} + Q(w) \end{array}$$

est une forme quadratique sur G : comme $\dim(G) = n-2$, par hypothèse de récurrence, il existe $n-2$ formes linéaires $\rho_3, \ldots, \rho_n \in \mathcal{L}(G, \mathbb{K})$ sur G et des scalaires $\alpha_3, \ldots, \alpha_n \in \mathbb{K}$ tels que, pour tout $w \in G$,

$$-2\frac{f_1(w)f_2(w)}{a_{1,2}} + Q(w) = \sum_{k=3}^{n} \alpha_k \big(\rho_k(w)\big)^2.$$

Si, pour tout $k \in \{3, \ldots, n\}$, on note

$$l_k : \begin{array}{ccc} E & \to & \mathbb{K} \\ v = \sum_{k=1}^{n} x_k e_k & \mapsto & \rho_k\left(\sum_{k=3}^{n} x_k e_k\right) \end{array},$$

on obtient alors l'égalité

$$q(v) = \frac{a_{1,2}}{2}\big(l_1(v)\big)^2 - \frac{a_{1,2}}{2}\big(l_2(v)\big)^2 + \sum_{k=3}^{n} \alpha_k \big(l_k(w)\big)^2.$$

De plus, la famille $\{l_1, l_2, \ldots, l_n\}$ de $\mathcal{L}(E, \mathbb{K})$ est une famille libre : contrairement à l_1 et l_2, les expressions des formes linéaires de la famille libre $\{l_3, \ldots, l_3\}$ ne dépendent pas des coordonnées x_1 et x_2, et les formes linéaires l_1 et l_2 sont linéairement indépendantes.

□

Exemple 4.6.2. 1. On considère la forme quadratique

$$q : \begin{array}{ccc} \mathbb{R}^3 & \to & \mathbb{R} \\ (x, y, z) & \mapsto & 2x^2 + y^2 - z^2 + 3xy - 5xz + 6yz \end{array}$$

sur $\mathbb{R}^3$. Pour tout $(x,y,z) \in \mathbb{R}^3$, on a

$$\begin{array}{rcl} q(x,y,z) & = & 2x^2+3xy-5xz+y^2-z^2+6yz \\ & = & 2\left(x^2+x\dfrac{3y-5z}{2}\right)+y^2-z^2+6yz \\ & = & 2\left(x+\dfrac{3y-5z}{4}\right)^2-\dfrac{(3y-5z)^2}{8}+y^2-z^2+6yz \\ & = & 2\left(x+\dfrac{3y-5z}{4}\right)^2+\dfrac{1}{8}\left(-y^2-33z^2+78yz\right) \\ & = & 2\left(x+\dfrac{3y-5z}{4}\right)^2+\dfrac{1}{8}\left[-\left(y^2-78yz\right)-33z^2\right] \\ & = & 2\left(x+\dfrac{3y-5z}{4}\right)^2+\dfrac{1}{8}\left[-(y-39z)^2+1521z^2-33z^2\right] \\ & = & 2\left(x+\dfrac{3y-5z}{4}\right)^2-\dfrac{1}{8}(y-39z)^2+186z^2 \end{array}$$

2. On considère la forme quadratique

$$q: \begin{array}{ccc} \mathbb{C}^3 & \to & \mathbb{C} \\ (x,y,z) & \mapsto & xy+ixz-(1-i)yz \end{array}$$

sur $\mathbb{C}^3$. Pour tout $(x,y,z) \in \mathbb{C}^3$, on a

$$\begin{array}{rcl} q(x,y,z) & = & xy+ixz-(1-i)yz \\ & = & \big(x-(1-i)z\big)\big(y+iz\big)+(1-i)iz^2 \\ & = & \dfrac{1}{4}\left[\big(x+y+(-1+2i)z\big)^2-\big(x-y-z\big)^2\right]+(1+i)z^2 \\ & = & \dfrac{1}{4}\big(x+y+(-1+2i)z\big)^2-\dfrac{1}{4}\big(x-y-z\big)^2+(1+i)z^2 \end{array}$$

Remarque 4.6.3. Si $q=\sum_{k=1}^{n}\alpha_k(l_k)^2$ avec $\alpha_1,\dots,\alpha_n \in \mathbb{K}$ et $l_1,\dots,l_n \in \mathcal{L}(E,\mathbb{K})$, alors, pour tout $v,w \in E$,

$$\varphi(v,w)=\sum_{k=1}^{m}\alpha_k l_k(v) l_k(w)$$

(cf. exemple 4.3.3 2.)

Nous allons utiliser la méthode de réduction de Gauss précédente pour déterminer une base orthogonale relativement à φ :

Théorème 4.6.4. *Soient $\alpha_1,\dots,\alpha_n \in \mathbb{K}$ et soient $l_1,\dots,l_n \in \mathcal{L}(E,\mathbb{K})$ des formes linéaires linéairement indépendantes tels que $q=\sum_{k=1}^{n}\alpha_k(l_k)^2$. Il existe une base orthogonale $\mathcal{B}=\{e_1,\dots,e_n\}$*

de E relativement à φ telle que pour tout $k \in \{1, \ldots, n\}$, $q(e_k) = \alpha_k$. Ainsi,

$$\mathrm{Mat}_{\mathcal{B}}^{\mathrm{fb}}(\varphi) = \begin{pmatrix} \alpha_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix}$$

Démonstration. Pour construire une base orthogonale avec les propriétés ci-dessus, on utilise la théorie de la dualité linéaire : on renvoie à l'annexe située à la fin de ce polycopié et consacrée à ce sujet.

Notons $\mathcal{B} = \{e_1, \ldots, e_n\}$ l'unique base de E telle que, pour tous $k, l \in \{1, \ldots, n\}$, $l_k(e_l) = \delta_{k,l}$ ($\{e_1, \ldots, e_n\}$ est la base dite antéduale de la base $\{l_1, \ldots, l_n\}$ de $\mathcal{L}(E, \mathbb{K})$: cf. corollaire et définition 5.4.2). On a alors, pour tous $r, s \in \{1, \ldots, n\}$,

$$\varphi(e_r, e_s) = \sum_{k=1}^{m} \alpha_k l_k(e_r) l_k(e_s) = \sum_{k=1}^{m} \alpha_k \delta_{k,r} \delta_{k,s} = \begin{cases} 0 & \text{si } r \neq s, \\ \alpha_r & \text{si } r = s. \end{cases}$$

□

Exemple 4.6.5. 1. On considère la forme quadratique de l'exemple 4.6.2 1. On a, pour tout $(x, y, z) \in \mathbb{R}^3$,

$$q(x, y, z) = 2x^2 + y^2 - z^2 + 3xy - 5xz + 6yz = 2\left(x + \frac{3y - 5z}{4}\right)^2 - \frac{1}{8}(y - 39z)^2 + 186z^2.$$

Notons $l_1 : \mathbb{R}^3 \to \mathbb{R}$; $(x, y, z) \mapsto x + \frac{3y-5z}{4}$, $l_2 : \mathbb{R}^3 \to \mathbb{R}$; $(x, y, z) \mapsto y - 39z$ et $l_3 : \mathbb{R}^3 \to \mathbb{R}$; $(x, y, z) \mapsto z$: la famille $\mathcal{C} := \{l_1, l_2, l_3\}$ est une base de $\mathcal{L}\left(\mathbb{R}^3, \mathbb{R}\right)$. Pour déterminer la base antéduale $\mathcal{B}$ de $\mathcal{C}$, on considère la matrice de passage

$$P_{\mathcal{B}_{\mathrm{can}}^* \to \mathcal{C}} = \begin{pmatrix} 1 & 0 & 0 \\ \frac{3}{4} & 1 & 0 \\ -\frac{5}{4} & -39 & 1 \end{pmatrix}$$

de la base duale de la base canonique $\mathcal{B}_{\mathrm{can}}$ de $\mathbb{R}^3$ à la base $\mathcal{C}$, et la matrice de passage de la base canonique de $\mathbb{R}^3$ à la base $\mathcal{B}$ est alors

$$P_{\mathcal{B}_{\mathrm{can}} \to \mathcal{B}} = {}^t P_{\mathcal{B}_{\mathrm{can}}^* \to \mathcal{C}}^{-1} = \begin{pmatrix} 1 & -\frac{3}{4} & -28 \\ 0 & 1 & 39 \\ 0 & 0 & 1 \end{pmatrix}$$

(cf. proposition 5.4.1 et corollaire et définition 5.4.2). La base antéduale de $\mathcal{C}$ est donc la famille

$$\mathcal{B} = \left\{(1, 0, 0), \left(-\frac{3}{4}, 1, 0\right), (-28, 39, 1)\right\},$$

qui, par la preuve du théorème précédent, est une base orthogonale relativement à la forme polaire φ de q, et on a

$$\mathrm{Mat}_{\mathcal{B}}^{\mathrm{fb}}(\varphi) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -\frac{1}{8} & 0 \\ 0 & 0 & 186 \end{pmatrix}.$$

2. On considère la forme quadratique de l'exemple 4.6.2 2. On a, pour tout $(x, y, z) \in \mathbb{C}^3$,

$$q(x,y,z) = xy + ixz - (1-i)yz = \frac{1}{4}\big(x + y + (-1+2i)z\big)^2 - \frac{1}{4}\big(x-y-z\big)^2 + (1+i)z^2.$$

Notons $l_1 : \mathbb{C}^3 \to \mathbb{C}$; $(x,y,z) \mapsto x + y + (-1+2i)z$, $l_2 : \mathbb{C}^3 \to \mathbb{C}$; $(x,y,z) \mapsto x - y - z$ et $l_3 : \mathbb{C}^3 \to \mathbb{C}$; $(x,y,z) \mapsto z$: la famille $\mathcal{C} := \{l_1, l_2, l_3\}$ est une base de $\mathcal{L}\left(\mathbb{C}^3, \mathbb{C}\right)$. On considère alors la matrice de passage

$$P_{\mathcal{B}^*_{\text{can}} \to \mathcal{C}} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ -1+2i & -1 & 1 \end{pmatrix}$$

de la base duale de la base canonique $\mathcal{B}_{\text{can}}$ de $\mathbb{C}^3$ à la base $\mathcal{C}$, et la matrice de passage de la base canonique de $\mathbb{C}^3$ à la base antéduale $\mathcal{B}$ de $\mathcal{C}$ est alors

$$P_{\mathcal{B}_{\text{can}} \to \mathcal{B}} = {}^tP^{-1}_{\mathcal{B}^*_{\text{can}} \to \mathcal{C}} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 1-i \\ \frac{1}{2} & -\frac{1}{2} & -i \\ 0 & 0 & 1 \end{pmatrix}.$$

La base

$$\mathcal{B} = \left\{ \left(\frac{1}{2}, \frac{1}{2}, 0\right), \left(\frac{1}{2}, -\frac{1}{2}, 0\right), (1-i, -i, 1) \right\}$$

est donc une base de $\mathbb{C}^3$ orthogonale relativement à la forme polaire φ de q, et on a

$$\mathrm{Mat}^{\mathrm{fb}}_{\mathcal{B}}(\varphi) = \begin{pmatrix} \frac{1}{4} & 0 & 0 \\ 0 & -\frac{1}{4} & 0 \\ 0 & 0 & 1+i \end{pmatrix}.$$

Remarque 4.6.6. Si, dans l'exemple 4.6.5 1., on considère la base

$$\mathcal{B}' = \left\{ \frac{1}{\sqrt{2}}(1,0,0), \sqrt{8}\left(-\frac{3}{4}, 1, 0\right), \frac{1}{\sqrt{186}}(-28, 39, 1) \right\}$$

de $\mathbb{R}^3$, on a

$$\mathrm{Mat}^{\mathrm{fb}}_{\mathcal{B}'}(\varphi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Si, dans l'exemple 4.6.5 2., on considère la base

$$\mathcal{B} = \left\{ 2\left(\frac{1}{2}, \frac{1}{2}, 0\right), 2i\left(\frac{1}{2}, -\frac{1}{2}, 0\right), \sqrt[4]{2}\, e^{i\frac{\pi}{8}}(1-i, -i, 1) \right\}$$

de $\mathbb{C}^3$, on a

$$\mathrm{Mat}^{\mathrm{fb}}_{\mathcal{B}'}(\varphi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

4.7 Signature d'une forme quadratique réelle

Dans cette section, on suppose que E un $\mathbb{R}$-espace vectoriel de dimension finie $n \in \mathbb{N}\backslash\{0\}$. Soit φ une forme bilinéaire symétrique sur E de forme quadratique q.

Définition 4.7.1. *On dit que φ (ou q) est positive, resp. négative, si pour tout $v \in E$, $q(v) = \varphi(v,v) \geqslant 0$, resp. $q(v) = \varphi(v,v) \leqslant 0$.*

Exemple 4.7.2.
1. La forme quadratique $q_1 : \mathbb{R}^2 \times \mathbb{R}^2$; $(x,y) \mapsto x^2 + y^2$ est positive.
2. La forme quadratique $q_2 : \mathbb{R}^2 \times \mathbb{R}^2$; $(x,y) \mapsto -x^2 - y^2$ est négative.
3. La forme quadratique $q_3 : \mathbb{R}^2 \times \mathbb{R}^2$; $(x,y) \mapsto x^2 - y^2$ n'est ni positive ni négative ($q(1,0) = 1$ et $q(0,1) = -1$).

On avait vu à la proposition 4.4.12 que si φ est définie, alors φ est non dégénérée. Sur $\mathbb{R}$, on peut compléter cette assertion :

Proposition 4.7.3. *La forme bilinéaire symétrique φ est définie ssi φ est non dégénérée et φ est positive ou négative.*

Démonstration. On montre le sens direct par contraposée :

- Si φ n'est ni positive ni négative, il existe $v, w \in E$ tels que $q(v) < 0$ et $q(w) > 0$. On applique alors le théorème des valeurs intermédiaires à la fonction continue

$$f : \begin{array}{ccc} [0;1] & \to & \mathbb{R} \\ t & \mapsto & q(v + t(v-w)) \end{array} :$$

on a $f(0) = q(v) < 0$ et $f(1) = q(w) > 0$ donc il existe $t \in [0;1]$ tel que $f(t) = q(v + t(v-w)) = 0$. De plus $v + t(v-w) \neq 0_E$ car la famille $\{v, w\}$ est libre (si elle était liée, les quantités $q(v)$ et $q(w)$ seraient de même signe ou toutes deux nulles). La forme bilinéaire symétrique φ est donc non définie.

- Si φ est dégénérée, alors elle est non définie par la proposition 4.4.12.

Montrons maintenant la réciproque : supposons que φ est positive, resp. négative, et que φ est non dégénérée. Comme φ est définie si et seulement si $-\varphi$ est définie, on peut supposer sans perdre de généralité que φ est positive. Pour montrer que φ est définie, considérons un vecteur v de E tel que $\varphi(v,v) = q(v) = 0$ et supposons par l'absurde que $v \neq 0_E$. Soit $w \in E\backslash\{0_E\}$ tel que $\varphi(v,w) \neq 0$ (φ est non dégénérée donc v, supposé non nul, ne peut être orthogonal à tous les vecteurs de E) : en multipliant éventuellement w par -1, on peut supposer que la quantité $\varphi(v,w)$ est strictement positive. Pour $t \in]0; +\infty[$, on a alors

$$0 \leqslant q(v - tw) = q(v) - 2t\varphi(v,w) + t^2 q(w) = -2t\varphi(v,w) + t^2 q(w).$$

donc $0 \leqslant -2\varphi(v,w) + tq(w)$ (car $t > 0$). Or cette dernière quantité peut être rendue strictement inférieur à 0 pour t suffisamment petit (car $\varphi(v,w) > 0$ et $q(w) > 0$), d'où une contradiction : on en conclut que $v = 0_E$. □

Pour montrer le théorème qui suit, nous aurons besoin du fait suivant :

Lemme 4.7.4. *Soit $\{v_1, \dots, v_p\}$ une famille libre et φ-orthogonale de E telle que pour tout $k \in \{1, \dots, p\}$, $q(v_k) > 0$. Alors, pour tout $v \in \mathrm{Vect}\{v_1, \dots, v_p\}$, $q(v) \geqslant 0$, et, si $v \neq 0_E$, $q(v) > 0$.*

Démonstration. On montre ce résultat par récurrence sur p. Si $p = 1$ et si $v = \lambda_1 v_1$ avec $\lambda_1 \in \mathbb{K}$, alors $q(v) = q(\lambda_1 v_1) = \lambda_1^2 q(v_1) \geqslant 0$, car $q(v_1) > 0$, et $q(v) > 0$ si $\lambda_1 \neq 0$.

Supposons maintenant le résultat vérifié au rang $p-1$ pour $p \in \mathbb{N}\backslash\{0; 1\}$ fixé. Si $v = \sum_{k=1}^{p} \lambda_k v_k$ avec $\lambda_1, \dots, \lambda_p \in \mathbb{K}$, on a

$$\begin{aligned}
q\left(\sum_{k=1}^{p} \lambda_k v_k\right) &= q\left(\lambda_p v_p + \sum_{k=1}^{p-1} \lambda_k v_k\right) \\
&= \lambda_p^2\, q(v_p) + 2\lambda_p\, \varphi\left(v_p, \sum_{k=1}^{p-1} \lambda_k v_k\right) + q\left(\sum_{k=1}^{p-1} \lambda_k v_k\right) \\
&= \lambda_p^2\, q(v_p) + 2\lambda_p \sum_{k=1}^{p-1} \lambda_k\, \varphi(v_p, v_k) + q\left(\sum_{k=1}^{p-1} \lambda_k v_k\right) \\
&= \lambda_p^2\, q(v_p) + q\left(\sum_{k=1}^{p-1} \lambda_k v_k\right).
\end{aligned}$$

Cette dernière quantité est positive car $q(v_p) > 0$ et $q\left(\sum_{k=1}^{p-1} \lambda_k v_k\right) \geqslant 0$ par hypothèse de récurrence. Elle est de plus strictement positive si $v \neq 0_E$ i.e. si les scalaires $\lambda_1, \dots, \lambda_p$ sont non tous nuls (la famille $\{v_1, \dots, v_p\}$ est libre), car

- si $\lambda_p \neq 0$, $\lambda_p^2\, q(v_p) > 0$,
- si $\lambda_p = 0$, $\sum_{k=1}^{p-1} \lambda_k v_k = v \neq 0_E$ donc, par hypothèse de récurrence, $q\left(\sum_{k=1}^{p-1} \lambda_k v_k\right) > 0$.

□

Remarque 4.7.5. Si $\{v_1, \dots, v_p\}$ est une famille libre et φ-orthogonale de E telle que pour tout $k \in \{1, \dots, p\}$, $q(v_k) < 0 \Leftrightarrow (-q)(e_k) > 0$, alors pour tout $v \in \mathrm{Vect}\{v_1, \dots, v_p\}$, $(-q)(v) \geqslant 0 \Leftrightarrow q(v) \leqslant 0$, et, si $v \neq 0_E$, $(-q)(v) > 0 \Leftrightarrow q(v) < 0$.

Le théorème ci-dessous définit la notion de "signature" d'une forme bilinéaire symétrique sur $\mathbb{R}$:

Théorème et Définition 4.7.6. *Soit $\mathcal{B} = \{e_1, \dots, e_n\}$ une base φ-orthogonale de E. On note r, resp. s, le nombre de coefficients strictement positifs, resp. strictement négatifs, de la matrice diagonale $\mathrm{Mat}_{\mathcal{B}}^{\mathrm{fb}}(\varphi)$. Les nombres r et s sont indépendants de la base orthogonale $\mathcal{B}$ considérée et on appelle signature de φ (ou de q) le couple (r, s).*

Démonstration. Soit $\mathcal{B}' = \{e'_1, \dots, e'_n\}$ une autre base φ-orthogonale de E et notons r', resp. s', le nombre de coefficients strictement positifs, resp. strictement négatifs, de la matrice diagonale $\mathrm{Mat}^{\mathrm{fb}}_{\mathcal{B}'}(\varphi)$.

Quitte à permuter les vecteurs de la base $\mathcal{B}$, resp. $\mathcal{B}'$, on peut supposer que pour tout $k \in \{1, \dots, r\}$, $q(e_k) > 0$, resp. pour tout $k \in \{1, \dots, r'\}$, $q(e'_k) > 0$, et que pour tout $l \in \{r+1, \dots, n\}$, $q(e_l) \leqslant 0$, resp. pour tout $l \in \{r'+1, \dots, n\}$, $q(e'_l) \leqslant 0$.

Notons ensuite $F := \mathrm{Vect}\{e_1, \dots, e_r\}$ et $G' := \mathrm{Vect}\{e'_{r'+1}, \dots, e'_n\}$. On a $F \cap G' = \{0_E\}$: si $v \in F \cap G'$ et $v \neq 0_E$ alors $q(v) > 0$ et $q(v) \leqslant 0$ par le lemme 4.7.4 et la remarque 4.7.5, ce qui est impossible.

Ainsi, $\dim(F) + \dim(G') = \dim(F \oplus G') \leqslant n$ (car $F \oplus G'$ est un sous-espace vectoriel de E). Or $\dim(F) = r$ et $\dim(G') = n - r'$ donc $r - r' \leqslant 0$. En échangeant les rôles de $\mathcal{B}$ et $\mathcal{B}'$, on obtient ensuite que $r' - r \leqslant 0$ et donc finalement $r = r'$.

Enfin, comme $r + s = \mathrm{rg}_\varphi = r' + s'$ (par la proposition 4.5.6), on a également $s = s'$. □

Exemple 4.7.7. La signature de la forme bilinéaire symétrique de l'exemple 4.6.5 1. est $(2, 1)$

Notons (r, s) la signature de la forme bilinéaire symétrique φ.

Proposition 4.7.8.
- *φ est positive ssi $s = 0$,*
- *φ est négative ssi $r = 0$,*
- *φ est non dégénérée ssi $r + s = n$.*

Démonstration. Soit $\mathcal{B} = \{e_1, \dots, e_n\}$ une base φ-orthogonale de E et écrivons

$$\mathrm{Mat}^{\mathrm{fb}}_{\mathcal{B}}(\varphi) = \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix}$$

- Si φ est positive alors, pour tout $k \in \{1, \dots, n\}$, $a_k = q(e_k) \geqslant 0$ i.e. aucun des coefficients diagonaux de $\mathrm{Mat}^{\mathrm{fb}}_{\mathcal{B}}(\varphi)$ n'est strictement négatif i.e. $s = 0$.

  Réciproquement, si $s = 0$, on a, pour tout $k \in \{1, \dots, n\}$, $a_k \geqslant 0$ et donc, pour tout $v \in E$, $q(v) \geqslant 0$.

- La signature de la forme bilinéaire symétrique $-\varphi$ est le couple (s, r), et φ est négative ssi $-\varphi$ est positive ssi $r = 0$.

- Enfin, φ est non dégénérée ssi $n = \mathrm{rg}_\varphi = r + s$.

□

Chapitre 5

Annexe : Dualité linéaire

5.1 Introduction

La dualité linéaire est la théorie des formes linéaires sur un espace vectoriel, c'est-à-dire, pour $\mathbb{K}$ un corps commutatif quelconque, la théorie des applications linéaires $E \to \mathbb{K}$ où E est un espace vectoriel sur $\mathbb{K}$. On peut également voir la dualité linéaire comme la théorie des équations linéaires sur un espace vectoriel. En particulier, cette théorie nous donne une correspondance explicite entre les sous-espaces vectoriels d'un espace vectoriel E et les systèmes d'équations linéaires sur E. La dualité linéaire nous fournit également une interprétation vectorielle de l'opération de transposition sur les matrices.

Tout au long de ce chapitre, $\mathbb{K}$ désigne un corps commutatif quelconque.

5.2 Formes linéaires sur un espace vectoriel et espace dual

Soit E un espace vectoriel sur $\mathbb{K}$.

Définition 5.2.1. *On appelle forme linéaire sur E toute application linéaire de E dans $\mathbb{K}$. L'ensemble des formes linéaires sur E est appelé espace dual de E et noté E^*.*

Exemple 5.2.2 (exemple “fil rouge”). L'application $\varphi : \begin{array}{ccc} \mathbb{R}^3 & \to & \mathbb{R} \\ (x,y,z) & \mapsto & 2x+3y-5z \end{array}$ est une forme linéaire sur $\mathbb{R}^3$.

Remarque 5.2.3.

- $E^* = \mathcal{L}(E,\mathbb{K})$ est un espace vectoriel sur $\mathbb{K}$.
- Si E est de dimension finie $n \in \mathbb{N}\backslash\{0\}$ et si $\mathcal{B} = \{e_1, \ldots, e_n\}$ est une base de E, alors, pour tout vecteur v de E de coordonnées $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ dans la base $\mathcal{B}$ et toute forme linéaire φ de

E^*, on a

$$\begin{aligned}\varphi(v) &= \varphi(x_1e_1 + \cdots + x_ne_n)\\ &= x_1 \underbrace{\varphi(e_1)}_{\in\mathbb{K}} + \cdots + x_n \underbrace{\varphi(e_n)}_{\in\mathbb{K}}\\ &= (\varphi(e_1) \cdots \varphi(e_n)) \begin{pmatrix} x_1\\ \vdots \\ x_n\end{pmatrix}\end{aligned}$$

Exemple 5.2.4 (suite de l'exemple "fil rouge"). Pour tout vecteur (x, y, z) de $\mathbb{R}^3$, on a

$$\varphi(x, y, z) = (2 \;\; 3 \;\; -5) \begin{pmatrix} x\\ y\\ z\end{pmatrix}$$

Définition 5.2.5. *Soit $\varphi \in E^*$ une forme linéaire sur E non identiquement nulle. On appelle <u>hyperplan de E déterminé par φ</u> le sous-espace vectoriel* Ker φ *de E.*

Exemple 5.2.6 (suite de l'exemple "fil rouge"). L'hyperplan de $\mathbb{R}^3$ déterminé par φ est le sous-espace vectoriel $\{(x, y, z) \in \mathbb{R}^3 \mid 2x + 3y - 5y = 0\}$ de $\mathbb{R}^3$.

L'appellation "hyperplan" est justifiée par le fait que, si E est de dimension finie $n \in \mathbb{N}\backslash\{0\}$, l'hyperplan déterminée par une forme linéaire sur E non identiquement nulle est effectivement un sous-espace vectoriel de E de dimension $n - 1$. Nous allons montrer ce fait ci-dessous, ainsi que sa réciproque :

Proposition 5.2.7. *Supposons que E est de dimension finie $n \in \mathbb{N}\backslash\{0\}$.*

Les hyperplans de E déterminés par les formes linéaires non nulles de E^* sont exactement les hyperplans linéaires de E i.e. ("id est" : c'est-à-dire) les sous-espaces vectoriels de E de dimension $n - 1$.

Démonstration. Soit $\varphi \in E^*\backslash\{0\}$. L'image Im φ de φ étant un sous-espace vectoriel de $\mathbb{K}$, la dimension de Im φ (sur $\mathbb{K}$) est inférieure ou égale à 1 (la dimension de $\mathbb{K}$ sur $\mathbb{K}$ est 1). Comme φ est non identiquement nulle, la dimension de Im φ ne peut être 0. La dimension de Im φ est donc 1 et, par le théorème du rang,

$$\dim(\text{Ker } \varphi) = \dim(E) - \dim(\text{Im } \varphi) = n - 1,$$

i.e. l'hyperplan de E déterminé par φ est de dimension $n - 1$.

Réciproquement, soit H un sous-espace vectoriel de E de dimension $n - 1$. Soit v_0 un vecteur de E n'appartenant pas à H. Alors les sous-espaces vectoriels H et $\text{Vect}\{v_0\}$ de E sont en somme directe et, par un argument de dimension, $E = H \oplus \text{Vect}\{v_0\}$. Ainsi, tout vecteur v de E se décompose de façon unique en une somme $v = u_v + \lambda_v v_0$ avec $u_v \in H$ et $\lambda_v \in \mathbb{K}$. Si φ désigne alors la forme linéaire $\begin{array}{ccc} E & \to & \mathbb{K}\\ v & \mapsto & \lambda_v\end{array}$, on a $H = \text{Ker } \varphi$, i.e. H est l'hyperplan de E déterminé par la forme linéaire φ. □

Remarque 5.2.8. Reprenons les notations de la remarque 5.2.3. Si $\varphi \in E^*$, l'hyperplan de E déterminé par φ est le sous-espace vectoriel de E caractérisé par l'équation linéaire

$$\underbrace{\varphi(e_1)}_{\in\mathbb{K}} x_1 + \cdots + \underbrace{\varphi(e_n)}_{\in\mathbb{K}} x_n = 0$$

en les coordonnées $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ dans la base $\mathcal{B}$.

Réciproquement, à tout sous-espace vectoriel H de E caractérisé par une équation linéaire

$$a_1x_1 + \cdots + a_nx_n = 0$$

avec $(a_1, \ldots, a_n) \in \mathbb{K}^n \backslash \{(0, \ldots, 0)\}$, on peut associer la forme linéaire

$$\varphi : \begin{array}{ccc} E & \to & \mathbb{K} \\ x_1e_1 + \cdots + x_ne_n & \mapsto & a_1x_1 + \cdots + a_nx_n \end{array},$$

et alors $H = \text{Ker }\varphi$.

On obtient ainsi une "correspondance" entre l'espace dual E^* de E et les équations linéaires en les coordonnées dans la base $\mathcal{B}$. A noter que l'espace des solutions d'un système d'équations linéaires peut être vu comme une intersection d'hyperplans.

5.3 Base duale

Soit E un espace vectoriel sur $\mathbb{K}$ de dimension finie $n \in \mathbb{N}\backslash\{0\}$.

En utilisant le fait que $\dim(\mathcal{L}(E, \mathbb{K})) = \dim(E) \times \dim(\mathbb{K}) = \dim(E)$, on peut montrer directement que $\dim(E^*) = \dim(E)$. On peut également le montrer en associant à toute base de E une base de E^* :

Théorème et Définition 5.3.1. *Soit $\mathcal{B} = \{e_1, \ldots, e_n\}$ une base de E.*

Pour $i \in \{1, \ldots, n\}$, on note e_i^* la forme linéaire sur E définie par

$$\textit{pour tout } j \in \{1, \ldots, n\},\ e_i^*(e_j) = \delta_{ij} = \begin{cases} 1 & \textit{ăsi } i = j, \\ 0 & \textit{ăsi } i \neq j. \end{cases}$$

La famille $\{e_1^*, \ldots, e_n^*\}$ de E^* est une base de E^*, appelée base duale de $\mathcal{B}$. On la note $\mathcal{B}^*$.

Démonstration. Soit $i \in \{1, \ldots, n\}$. Commençons par remarquer que, par définition, si v est un vecteur de E de coordonnées $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ dans la base $\mathcal{B}$,

$$e_i^*(v) = e^*(x_1e_1 + \cdots + x_ne_n) = x_1e_i^*(e_1) + \ldots + x_ne_i^*(e_n) = x_i,$$

autrement dit e_i^* associe à tout vecteur v de E sa $i^{\text{ème}}$ coordonnée dans la base $\mathcal{B}$.

À présent, montrons que la famille $\{e_1^*, \ldots, e_n^*\}$ de E^* est libre : soient $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ tels que $\lambda_1 e_1^* + \ldots + \lambda_n e_n^*$ soit la forme linéaire nulle, i.e., pour tout vecteur v de E, $\lambda_1 e_1^*(v) + \ldots + \lambda_n e_n^*(v) = 0$. En particulier, pour tout $j \in \{1, \ldots, n\}$, $0 = \lambda_1 e_1^*(e_j) + \ldots + \lambda_n e_n^*(e_j) = \lambda_j$ et la famille $\{e_1^*, \ldots, e_n^*\}$ de E^* est donc libre.

Montrons ensuite que la famille $\{e_1^*, \ldots, e_n^*\}$ engendre E^*. Soit donc $\varphi \in E^*$, et soit v un vecteur de coordonnées $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ dans la base $\mathcal{B}$. On a alors

$$\varphi(v) = \varphi(x_1 e_1 + \cdots + x_n e_n) = \underbrace{\varphi(e_1)}_{\in \mathbb{K}} x_1 + \cdots + \underbrace{\varphi(e_n)}_{\in \mathbb{K}} x_n = \varphi(e_1) e_1^*(v) + \cdots + \varphi(e_n) e_n^*(v).$$

Ainsi, $\varphi = \varphi(e_1) e_1^* + \cdots + \varphi(e_n) e_n^* \in \mathrm{Vect}\{e_1^*, \ldots, e_n^*\}$ et la famille $\{e_1^*, \ldots, e_n^*\}$ est donc génératrice de E^*. □

Remarque 5.3.2.

- On aurait pu se contenter de montrer le caractère libre ou le caractère générateur de la famille $\{e_1^*, \ldots, e_n^*\}$ puis d'utiliser le fait, établi précédemment, que $\dim(E^*) = \dim(E) = n$ pour montrer que la famille $\{e_1^*, \ldots, e_n^*\}$ est une base de E^*. On a cependant fait le choix de la démonstration "complète" ci-dessus pour son intérêt didactique.

- E et E^* étant deux espaces vectoriels de même dimension finie, ils sont isomorphes. Cependant, en général, ils ne le sont pas de façon "canonique" : un isomorphisme entre ces deux espaces vectoriels dépend, en général, d'un choix de bases pour E et E^*.

Exemple 5.3.3. Si $\mathcal{B} = \{e_1, \ldots, e_n\}$ est la base canonique de $\mathbb{K}^n$, pour tout $i \in \{1, \ldots, n\}$, e_i^* est la forme linéaire

$$e_i^* : \begin{array}{ccc} \mathbb{K}^n & \to & \mathbb{K} \\ (x_1, \ldots, x_n) & \mapsto & x_i \end{array}$$

Exemple 5.3.4. On considère la base $\mathcal{B} = \{e_1, e_2, e_3\}$ de $\mathbb{R}^3$ formé par les vecteurs $e_1 := (1,1,1)$, $e_2 := (1,0,-1)$ et $e_3 := (0,1,1)$. Déterminons la base duale $\mathcal{B}^*$ de $\mathcal{B}$: précisément, nous allons déterminer les expressions des formes linéaires e_1^*, e_2^* et e_3^* sur $\mathbb{R}^3$.

On cherche $a, b, c \in \mathbb{R}$ tels que, pour tout $(x_1, x_2, x_3) \in \mathbb{R}^3$, $e_1^*(x_1, x_2, x_3) = ax_1 + bx_2 + cx_3$. Or

$$\begin{cases} e_1^*(e_1) = 1 \\ e_1^*(e_2) = 0 \\ e_1^*(e_3) = 0 \end{cases} \Leftrightarrow \begin{cases} a + b + c = 1 \\ a \quad\quad - c = 0 \\ \quad\quad b + c = 0 \end{cases} \Leftrightarrow \begin{cases} a = 1 \\ b = -1 \\ c = 1 \end{cases}$$

Ainsi, e_1^* est l'application

$$e_1^* : \begin{array}{ccc} \mathbb{R}^3 & \to & \mathbb{R} \\ (x_1, x_2, x_3) & \mapsto & x_1 - x_2 + x_3 \end{array}$$

On cherche à présent $a, b, c \in \mathbb{R}$ tels que, pour tout $(x_1, x_2, x_3) \in \mathbb{R}^3$, $e_2^*(x_1, x_2, x_3) = ax_1 + bx_2 + cx_3$. Or

$$\begin{cases} e_2^*(e_1) = 0 \\ e_2^*(e_2) = 1 \\ e_2^*(e_3) = 0 \end{cases} \Leftrightarrow \begin{cases} a + b + c = 0 \\ a \quad\quad - c = 1 \\ \quad\quad b + c = 0 \end{cases} \Leftrightarrow \begin{cases} a = 0 \\ b = 1 \\ c = -1 \end{cases}$$

Ainsi, e_2^* est l'application

$$e_2^* : \begin{array}{ccc} \mathbb{R}^3 & \to & \mathbb{R} \\ (x_1, x_2, x_3) & \mapsto & x_2 - x_3 \end{array}$$

Enfin, on cherche $a, b, c \in \mathbb{R}$ tels que, pour tout $(x_1, x_2, x_3) \in \mathbb{R}^3$, $e_3^*(x_1, x_2, x_3) = ax_1 + bx_2 + cx_3$. Or

$$\begin{cases} e_3^*(e_1) = 0 \\ e_3^*(e_2) = 0 \\ e_3^*(e_3) = 1 \end{cases} \Leftrightarrow \begin{cases} a + b + c = 0 \\ a \quad\quad - c = 0 \\ \quad\quad b + c = 1 \end{cases} \Leftrightarrow \begin{cases} a = -1 \\ b = 2 \\ c = -1 \end{cases}$$

Ainsi, e_3^* est l'application

$$e_3^* : \begin{array}{ccc} \mathbb{R}^3 & \to & \mathbb{R} \\ (x_1, x_2, x_3) & \mapsto & -x_1 + 2x_2 - x_3 \end{array}$$

Remarque 5.3.5. Attention : parler de "dual d'un vecteur" n'a pas de sens. Si $\mathcal{B}_1$ et $\mathcal{B}_2$ sont deux bases de E et v est un vecteur de E appartenant à chacune de ces deux bases, les vecteurs "v^*" dans $\mathcal{B}_1^*$ et "v^*" dans $\mathcal{B}_2^*$ sont a priori différents (on devrait écrire $v^{*\mathcal{B}_1}$, respectivement $v^{*\mathcal{B}_2}$).

Reprenons les vecteurs $e_1 = (1, 1, 1)$ et $e_2 = (1, 0, -1)$ de $\mathbb{R}^3$ de l'exemple précédent mais posons cette fois $v_3 := (1, 0, 0)$. La famille $\mathcal{B}' := \{e_1, e_2, v_3\}$ est également une base de $\mathbb{R}^3$. Déterminons les formes linéaires de la base duale $\mathcal{B}'^*$ de $\mathcal{B}'$: on cherche $a, b, c \in \mathbb{R}$ tels que, pour tout $(x_1, x_2, x_3) \in \mathbb{R}^3$, $e_1^*(x_1, x_2, x_3) = ax_1 + bx_2 + cx_3$. Or

$$\begin{cases} e_1^*(e_1) = 1 \\ e_1^*(e_2) = 0 \\ e_1^*(v_3) = 0 \end{cases} \Leftrightarrow \begin{cases} a + b + c = 1 \\ a \quad\quad - c = 0 \\ a \quad\quad\quad = 0 \end{cases} \Leftrightarrow \begin{cases} a = 0 \\ b = 1 \\ c = 0 \end{cases}$$

Ainsi, e_1^* est l'application

$$e_1^* : \begin{array}{ccc} \mathbb{R}^3 & \to & \mathbb{R} \\ (x_1, x_2, x_3) & \mapsto & x_2 \end{array}$$

On cherche ensuite $a, b, c \in \mathbb{R}$ tels que, pour tout $(x_1, x_2, x_3) \in \mathbb{R}^3$, $e_2^*(x_1, x_2, x_3) = ax_1 + bx_2 + cx_3$. Or

$$\begin{cases} e_2^*(e_1) = 0 \\ e_2^*(e_2) = 1 \\ e_2^*(v_3) = 0 \end{cases} \Leftrightarrow \begin{cases} a + b + c = 0 \\ a \quad\quad - c = 1 \\ a \quad\quad\quad = 0 \end{cases} \Leftrightarrow \begin{cases} a = 0 \\ b = 1 \\ c = -1 \end{cases}$$

Ainsi, e_2^* est l'application

$$e_2^* : \begin{array}{ccc} \mathbb{R}^3 & \to & \mathbb{R} \\ (x_1, x_2, x_3) & \mapsto & x_2 - x_3 \end{array}$$

Enfin, on cherche $a, b, c \in \mathbb{R}$ tels que, pour tout $(x_1, x_2, x_3) \in \mathbb{R}^3$, $v_3^*(x_1, x_2, x_3) = ax_1 + bx_2 + cx_3$. Or

$$\begin{cases} v_3^*(e_1) = 0 \\ v_3^*(e_2) = 0 \\ v_3^*(v_3) = 1 \end{cases} \Leftrightarrow \begin{cases} a + b + c = 0 \\ a \qquad - c = 0 \\ a \qquad\quad = 1 \end{cases} \Leftrightarrow \begin{cases} a = 1 \\ b = -2 \\ c = 1 \end{cases}$$

Ainsi, v_3^* est l'application

$$v_3^* : \begin{array}{ccc} \mathbb{R}^3 & \to & \mathbb{R} \\ (x_1, x_2, x_3) & \mapsto & x_1 - 2x_2 + x_3 \end{array}$$

On remarque ainsi que $e_2^{*\mathcal{B}'} = e_2^{*\mathcal{B}}$ mais que $e_1^{*\mathcal{B}'} \neq e_1^{*\mathcal{B}}$. A noter également que, même si v_3 est le troisième vecteur de la base canonique de $\mathbb{R}^3$, $v_3^{*\mathcal{B}'}$ n'est pas l'application $\begin{array}{ccc} \mathbb{R}^3 & \to & \mathbb{R} \\ (x_1, x_2, x_3) & \mapsto & x_3 \end{array}$.

Dans la suite, $\mathcal{B} = \{e_1, \ldots, e_n\}$ désignera une base de E.

Proposition 5.3.6. *1. Pour toute forme linéaire $\varphi \in E^*$, $\varphi = \sum_{i=1}^{n} \varphi(e_i)e_i^*$, autrement dit φ a pour coordonnées $\begin{pmatrix} \varphi(e_1) \\ \vdots \\ \varphi(e_n) \end{pmatrix}$ dans la base duale $\mathcal{B}^*$ de $\mathcal{B}$.*

2. Pour tout vecteur $v \in E$, $v = \sum_{j=1}^{n} e_j^*(v)e_j$, autrement dit v a pour coordonnées $\begin{pmatrix} e_1^*(v) \\ \vdots \\ e_n^*(v) \end{pmatrix}$ dans la base $\mathcal{B}$.

Démonstration. 1. Soit $\varphi \in E^*$. Comme $\mathcal{B}^*$ est une base de E^*, il existe $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ (uniques) tels que $\varphi = \sum_{i=1}^{n} \lambda_i e_i^*$. Si $j \in \{1, \ldots, n\}$, on a alors $\varphi(e_j) = \sum_{i=1}^{n} \lambda_i e_i^*(e_j) = \lambda_j$, et donc $\varphi = \sum_{i=1}^{n} \varphi(e_i)e_i^*$.

2. Soit $v \in E$. Comme $\mathcal{B}$ est une base de E, il existe $\mu_1, \ldots, \mu_n \in \mathbb{K}$ (uniques) tels que $v = \sum_{j=1}^{n} \mu_j e_j$. Si $i \in \{1, \ldots, n\}$, on a alors $e_i^*(v) = \sum_{j=1}^{n} \mu_j e_i^*(e_j) = \mu_i$, et donc $v = \sum_{j=1}^{n} e_j^*(v)e_j$.

□

Remarque 5.3.7. Cela peut constituer un moyen "efficace" de déterminer les coordonnées d'une forme linéaire dans une base duale donnée, resp. ("respectivement") d'un vecteur dans une base donnée.

Exemple 5.3.8. • Reprenons la forme linéaire $\varphi : \begin{array}{ccc} \mathbb{R}^3 & \to & \mathbb{R} \\ (x_1, x_2, x_3) & \mapsto & 2x_1 + 3x_2 - 5x_3 \end{array}$ de l'exemple fil rouge et déterminons ses coordonnées dans la base duale $\mathcal{B}^*$ de l'exemple 5.3.4. On a $\varphi(e_1) = 0$, $\varphi(e_2) = 7$, $\varphi(e_3) = -2$ et donc $\varphi = 7e_2^* - 2e_3^*$. Remarquons que l'on n'a pas besoin de l'expression des formes linéaires de la base duale pour déterminer les coordonnées de φ dans celle-ci (les expressions obtenues dans l'exemple 5.3.4 nous permettent cependant de vérifier que la décomposition précédente est bien correcte).

- Si l'on considère le vecteur $v = (3, -4, 1)$ de $\mathbb{R}^3$, on obtient ses coordonnées dans la base $\mathcal{B}$ de l'exemple 5.3.4 en calculant $e_1^*(v) = 8$, $e_2^*(v) = -5$ et $e_3^*(v) = -12$. On a donc $v = 8e_1 - 5e_2 - 12e_3$.

La proposition 5.3.6 nous permet également de montrer de l'opération qui à toute base de E associe sa base duale est injective. Nous montrerons sa surjectivité dans la section suivante.

Corollaire 5.3.9. *ăSoit $\mathcal{B}' = \{f_1, \ldots, f_n\}$ une base de E. Si $\mathcal{B}^* = \mathcal{B}'^*$, alors $\mathcal{B} = \mathcal{B}'$.*

Démonstration. Soit $i \in \{1, \ldots, n\}$. D'après la proposition 5.3.6,

$$e_i = \sum_{j=1}^{n} f_j^*(e_i) f_j = \sum_{j=1}^{n} e_j^*(e_i) f_j = f_i.$$

Ainsi $\mathcal{B} = \mathcal{B}'$. □

Remarque 5.3.10. Comme, sur l'espace vectoriel de dimension finie E^*, on a accès à des bases, on peut utiliser les outils matriciels pour étudier les formes linéaires de E^*.

5.4 Aspects matriciels

Comme dans la section précédente, on considère un espace vectoriel E de dimension finie $n \in \mathbb{N}\backslash\{0\}$ et une base $\mathcal{B} = \{e_1, \ldots, e_n\}$ de E.

Commençons par une première remarque : si φ est une forme linéaire de E^* et v est un vecteur de E, alors, d'après la remarque 5.2.3,

$$\varphi(v) = {}^t\,\mathrm{Mat}_{\mathcal{B}*}(\varphi)\mathrm{Mat}_{\mathcal{B}}(v).$$

A présent, nous allons nous intéresser au changement de base pour les bases duales : précisément, soit $\mathcal{B}' = \{f_1, \ldots, f_n\}$ une autre base de E, on peut calculer la matrice de passage de la base duale $\mathcal{B}^*$ à la base duale $\mathcal{B}'^*$ à partir de la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$.

Proposition 5.4.1. *On a*

$$P_{\mathcal{B}^* \to \mathcal{B}'^*} = {}^tP_{\mathcal{B} \to \mathcal{B}'}{}^{-1} = {}^tP_{\mathcal{B}' \to \mathcal{B}}$$

(rappel : la transposition et l'inversion des matrices inversibles commutent).

Démonstration. Pour simplifier les écritures, on note $P = (p_{i\,j})_{1\leqslant i,j\leqslant n} := P_{\mathcal{B}\to\mathcal{B}'}$ et $Q = (q_{i\,j})_{1\leqslant i,j\leqslant n} := P_{\mathcal{B}^*\to\mathcal{B}'^*}$. Ainsi, pour $i, j \in \{1, \ldots, n\}$, $f_i^* = \sum_{k=1}^n q_{k\,i} e_k^*$ (il s'agit de la décomposition de f_i^* dans la base $\mathcal{B}^*$) et $f_j = \sum_{l=1}^n p_{l\,j} e_l$ (il s'agit de la décomposition de f_j dans la base $\mathcal{B}$). Nous allons montrer que $I_n = {}^tQP$ (et donc $Q = {}^tP^{-1}$).

Soient $i, j \in \{1, \ldots, n\}$. Le coefficient situé sur la ligne i et la colonne j de la matrice identité I_n est $\delta_{i,j}$ et, par définition de la base duale $\mathcal{B}'^* = \{f_1^*, \ldots, f_n^*\}$, on a

$$\begin{aligned} \delta_{i,j} = f_i^*(f_j) &= \left(\sum_{k=1}^n q_{k\,i} e_k^*\right)\left(\sum_{l=1}^n p_{l\,j} e_l\right) \\ &= \sum_{k=1}^n \sum_{l=1}^n q_{k\,i} p_{l\,j} e_k^*(e_l) \\ &= \sum_{k=1}^n \sum_{l=1}^n q_{k\,i} p_{l\,j} \delta_{k,l} \\ &= \sum_{k=1}^n q_{k\,i} p_{k\,j} \end{aligned}$$

Or $\displaystyle\sum_{k=1}^n q_{k\,i} p_{k\,j}$ est justement le coefficient situé sur la ligne i et la colonne j de la matrice produit ${}^tQP$. Ainsi, on a bien $I_n = {}^tQP$ i.e. $Q = {}^tP^{-1}$ i.e. $P_{\mathcal{B}^*\to\mathcal{B}'^*} = {}^tP_{\mathcal{B}\to\mathcal{B}'}{}^{-1}$. □

A l'aide de cette propriété, on peut également montrer la surjectivité, et donc la bijectivité (voir corollaire 5.3.9 ci-dessus), de l'opération qui associe à toute base $\mathcal{B}$ de E sa base duale $\mathcal{B}^*$ de E^* :

Corollaire et Définition 5.4.2. *Pour toute base $\mathcal{C}$ de E^*, il existe une et une seule base $\mathcal{B}$ de E telle que $\mathcal{C} = \mathcal{B}^*$. On appelle $\mathcal{B}$ la base antéduale de $\mathcal{C}$.*

Démonstration. Soit $\mathcal{C}$ une base de E^*. Fixons maintenant $\mathcal{B}_0$ une base quelconque de E et considérons la matrice $Q := {}^tP_{\mathcal{B}_0^*\to\mathcal{C}}{}^{-1}$. Comme il s'agit d'une matrice inversible de taille n, Q peut être considérée comme la matrice de passage $P_{\mathcal{B}_0\to\mathcal{B}}$ de la base $\mathcal{B}_0$ de E à une base $\mathcal{B}$ (les coordonnées des vecteurs de $\mathcal{B}$ dans la base $\mathcal{B}_0$ sont données par les colonnes de Q) et on a alors, d'après la proposition précédente,

$$P_{\mathcal{B}_0^*\to\mathcal{B}^*} = {}^tP_{\mathcal{B}_0\to\mathcal{B}}{}^{-1} = {}^tQ^{-1} = P_{\mathcal{B}_0^*\to\mathcal{C}},$$

de sorte que les coordonnées des vecteurs de la base $\mathcal{C}$ dans la base $\mathcal{B}_0^*$ sont les mêmes que les coordonnées des vecteurs de la base $\mathcal{B}^*$ dans la base $\mathcal{B}_0^*$ et donc $\mathcal{C} = \mathcal{B}^*$. □

Exemple 5.4.3. On considère les formes linéaires $\varphi_1 : \begin{array}{ccc} \mathbb{R}^3 & \to & \mathbb{R} \\ (x_1, x_2, x_3) & \to & x_1 + x_2 + x_3 \end{array}$, $\varphi_2 : \begin{array}{ccc} \mathbb{R}^3 & \to & \mathbb{R} \\ (x_1, x_2, x_3) & \to & -x_1 + x_3 \end{array}$ et $\varphi_3 : \begin{array}{ccc} \mathbb{R}^3 & \to & \mathbb{R} \\ (x_1, x_2, x_3) & \to & x_2 + x_3 \end{array}$ sur $\mathbb{R}^3$. La famille $\mathcal{C} = \{\varphi_1, \varphi_2, \varphi_3\}$ est une base de $\left(\mathbb{R}^3\right)^*$. Pour le voir, on écrit les coordonnées de φ_1, φ_2 et φ_3 dans la base duale

$\mathcal{B}_0^* = \{e_1^*, e_2^*, e_3^*\}$ de la base canonique $\mathcal{B}_0 = \{e_1, e_2, e_3\}$ de $\mathbb{R}^3$: on a $\varphi_1 = e_1^* + e_2^* + e_3^*$, $\varphi_2 = -e_1^* + e_3^*$ et $\varphi_3 = e_2^* + e_3^*$, et la matrice

$$P := \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

dont les colonnes sont les coordonnées de φ_1, φ_2 et φ_3 dans la base $\mathcal{B}^*$, est inversible.

On cherche maintenant à déterminer la base antéduale $\mathcal{B} = \{v_1, v_2, v_3\}$ de la base $\mathcal{C}$. On procède comme dans la démonstration précédente : la matrice P ci-dessus est la matrice de passage de $\mathcal{B}_0^*$ à $\mathcal{C}$ et la matrice de passage de la base $\mathcal{B}_0$ à la base $\mathcal{B}$ est alors la matrice ${}^tP^{-1}$. On obtient

$${}^tP^{-1} = \begin{pmatrix} 1 & 0 & -1 \\ -1 & -1 & 2 \\ 1 & 1 & -1 \end{pmatrix}$$

et on a donc $v_1 = e_1 - e_2 + e_3 = (1, -1, 1)$, $v_2 = -e_2 + e_3 = (0, -1, 1)$ et $v_3 = -e_1 + 2e_2 - e_3 = (-1, 2, -1)$.

5.5 Annulateur d'un sous-espace vectoriel et correspondance duale

Soit E un espace vectoriel de dimension finie $n \in \mathbb{N}\backslash\{0\}$.

Dans cette section, on va définir des outils de la dualité qui vont nous donner une correspondance explicite entre les sous-espaces vectoriels de E et les systèmes d'équations linéaires qui les caractérisent.

Soit F un sous-espace vectoriel de E et soit W un sous-espace vectoriel de E^*.

Définition 5.5.1.

- *L'ensemble, noté F^0, des formes linéaires de E^* qui s'annulent sur F est appelé annulateur de F.*
- *L'ensemble, noté W^0, des vecteurs de E qui sont annulés par toutes les formes linéaires de W est appelé annulateur de W.*

L'ensemble $F^0 = \{\varphi \in E^* \mid \text{ pour tout } v \in F,\ \varphi(v) = 0\}$ est un sous-espace vectoriel de E^* : la forme linéaire identiquement nulle sur E appartient à F^0 et, si $\varphi, \psi \in E^*$ et $\lambda, \mu \in \mathbb{K}$, pour tout $v \in E$, $(\lambda\varphi + \mu\psi)(v) = \lambda\varphi(v) + \mu\psi(v) = 0$.

De façon analogue, $W^0 = \{v \in E \mid \text{ pour tout } \varphi \in W,\ \varphi(v) = 0\}$ est un sous-espace vectoriel de E : le vecteur nul de E appartient à W^0 et, si $v, w \in E$ et $\lambda, \mu \in \mathbb{K}$, pour tout $\varphi \in E^*$, $\varphi(\lambda v + \mu w) = \lambda\varphi(v) + \mu\varphi(w) = 0$.

Proposition 5.5.2.

1. *Si $\{v_1, \ldots, v_p\}$ est une base de F, alors $F^0 = \{\varphi \in E^* \mid \varphi(v_1) = 0, \ldots, \varphi(v_p) = 0\}$.*
2. *Si $\{\varphi_1, \ldots, \varphi_q\}$ est une base de W, alors $W^0 = \{v \in E \mid \varphi_1(v) = 0, \ldots, \varphi_q(v) = 0\}$.*

Démonstration. 1. Soit $\varphi \in F^0$, alors, comme $v_1, \dots, v_p \in F$, $\varphi(v_1) = 0, \dots, \varphi(v_p) = 0$. Réciproquement, soit maintenant φ une forme linéaire sur E annulant les vecteurs $v_1, \dots, v_p$ et soit $v \in F$. Comme $\{v_1, \dots, v_p\}$ est une base de F, il existe $\lambda_1, \dots, \lambda_p \in \mathbb{K}$ (uniques) tels que $v = \lambda_1 v_1 + \cdots + \lambda_p v_p$ et alors

$$\varphi(v) = \lambda_1 \varphi(v_1) + \cdots + \lambda_p \varphi(v_p) = 0,$$

et φ appartient donc à F^0.

2. Soit $v \in W^0$, alors, comme $\varphi_1, \dots, \varphi_q \in W$, $\varphi_1(v) = 0, \dots, \varphi_q(v) = 0$. Réciproquement, soit maintenant v un vecteur de E annulé par les formes linéaires $\varphi_1, \dots, \varphi_q$ et soit $\varphi \in W$. Comme $\{\varphi_1, \dots, \varphi_q\}$ est une base de W, il existe $\mu_1, \dots, \mu_q \in \mathbb{K}$ (uniques) tels que $\varphi = \mu_1 \varphi_1 + \cdots + \mu_q \varphi_q$ et alors

$$\varphi(v) = \mu_1 \varphi_1(v) + \cdots + \mu_q \varphi_q(v) = 0,$$

et v appartient donc à W^0.

□

Remarquons que, si $\{\varphi_1, \dots, \varphi_q\}$ est une base de W, $W^0 = \bigcap_{i=1}^q \operatorname{Ker} \varphi_i$ et que, si $\mathcal{B} = \{e_1, \dots, e_n\}$ est une base de E, W^0 est le sous-espace vectoriel de E caractérisé par le système linéaire

$$\begin{cases} \varphi_1(e_1)x_1 + \cdots + \varphi_1(e_n)x_n = 0 \\ \qquad\qquad \vdots \\ \varphi_q(e_1)x_1 + \cdots + \varphi_q(e_n)x_n = 0 \end{cases}$$

en les coordonnées $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ dans la base $\mathcal{B}$.

La proposition suivante affirme notamment que si $W = F^0$, alors F peut être décrit par le système linéaire ci-dessus, autrement dit que les vecteurs de F sont exactement les solutions de ce système.

Proposition 5.5.3. *On a*

1. $\dim(E) = \dim(F) + \dim\left(F^0\right)$ *et* $\dim\left(E^*\right) = \dim(W) + \dim\left(W^0\right)$,

2. $\left(F^0\right)^0 = F$ *et* $\left(W^0\right)^0 = W$.

Démonstration. 1. Montrons tout d'abord que $\dim(E) = \dim(F) + \dim\left(F^0\right)$. Soit $\{v_1, \dots, v_p\}$ une base de F que l'on complète en une base $\mathcal{B} = \{v_1, \dots, v_p, v_{p+1}, \dots, v_n\}$ de E. Considérons la base duale $\mathcal{B}^* = \left\{v_1^*, \dots, v_p^*, v_{p+1}^*, \dots, v_n^*\right\}$ et montrons que la famille $\left\{v_{p+1}^*, \dots, v_n^*\right\}$ est une base de F^0 :

- Soit $i \in \{p+1, \dots, n\}$, alors $v_i^* \in F^0$ car, pour tout $j \in \{1, \dots, p\}$, $v_i^*(v_j) = 0$ $(i \neq j)$ et les vecteurs $v_1, \dots, v_p$ engendrent F.

- La famille $\{v_{p+1}^*, \ldots, v_n^*\}$ de E^* est libre comme sous-famille de la base $\mathcal{B}^*$.
- De plus, elle engendre F^0 : en effet, soit $\varphi \in F^0$, alors, d'après la proposition 5.3.6 i),

$$\begin{aligned}\varphi &= \varphi(v_1)v_1^* + \cdots + \varphi(v_p)v_p^* + \varphi(v_{p+1})v_{p+1}^* + \cdots + \varphi(v_n)v_n^* \\ &= \varphi(v_{p+1})v_{p+1}^* + \cdots + \varphi(v_n)v_n^* \in \operatorname{Vect}\{v_{p+1}^*, \ldots, v_n^*\}\end{aligned}$$

$(\varphi(v_1) = \cdots = \varphi(v_p) = 0$ car $v_1, \ldots, v_p \in F$ et $\varphi \in F^0)$.

En particulier, $\dim(F^0) = n - p = \dim(E) - \dim(F)$.

L'égalité $\dim(E^0) = \dim(W) + \dim(W^0)$ se démontre de façon tout à fait similaire, à l'aide de la notion de base antéduale (corollaire et définition 5.4.2) : soit $\{\varphi_1, \ldots, \varphi_q\}$ une base de W que l'on complète en une base $\mathcal{C} = \{\varphi_1, \ldots, \varphi_q, \varphi_{q+1}, \ldots, \varphi_n\}$ de E^* et notons $\mathcal{B} = \{v_1, \ldots, v_q, v_{q+1}, \ldots, v_n\}$ la base antéduale de $\mathcal{C}$. On montre que la famille $\{v_{q+1}, \ldots, v_n\}$ est une base de W^0 :

- Soit $j \in \{q+1, \ldots, n\}$, alors $v_j \in W^0$ car, pour tout $i \in \{1, \ldots, q\}$, $\varphi_i(v_j) = v_i^*(v_j) = 0$ $(i \neq j)$ et les vecteurs $\varphi_1, \ldots, \varphi_q$ engendrent W.
- La famille $\{v_{q+1}, \ldots, v_n\}$ de E est libre comme sous-famille de la base $\mathcal{B}$.
- De plus, elle engendre W^0 : en effet, soit $v \in W^0$, alors, d'après la proposition 5.3.6 ii),

$$\begin{aligned}v &= v_1^*(v)v_1 + \cdots + v_q^*(v)v_q + v_{q+1}^*(v)v_{q+1} + \cdots + v_n^*(v)v_n \\ &= \varphi_1(v)v_1 + \cdots + \varphi_q(v)v_q + \varphi_{q+1}(v)v_{q+1} + \cdots + \varphi_n(v)v_n \\ &= \varphi_{q+1}(v)v_{q+1} + \cdots + \varphi_n(v)v_n \in \operatorname{Vect}\{v_{q+1}, \ldots, v_n\}\end{aligned}$$

$(\varphi_1(v) = \cdots = \varphi_q(v) = 0$ car $\varphi_1, \ldots, \varphi_q \in W$ et $v \in W^0)$.

En particulier, $\dim(W^0) = n - q = \dim(E^*) - \dim(W) = \dim(E) - \dim(W)$.

2. Des deux égalités démontrées précédemment, on déduit la double inclusion $(F^0)^0 = F$: on a $F \subset (F^0)^0$ (car si $v \in F$ et $\varphi \in F^0$ alors $\varphi(v) = 0$) et

$$\dim\left((F^0)^0\right) = \dim(E) - \dim(F^0) = \dim(E) - (\dim(E) - \dim(F)) = \dim(F).$$

De même, $(W^0)^0 = W$ car $W \subset (W^0)^0$ (si $\varphi \in W$ et $v \in W^0$ alors $\varphi(v) = 0$) et

$$\dim\left((W^0)^0\right) = \dim(E) - \dim(W^0) = \dim(E) - (\dim(E) - \dim(W)) = \dim(W).$$

□

Cette proposition et sa démonstration nous donnent en particulier une méthode pour, à partir d'une base de F, obtenir un système d'équations linéaires "linéairement indépendantes" (i.e. d'équations correspondant à des formes linéaires linéairement indépendantes) décrivant F :

Exemple 5.5.4. Soit F le sous-espace vectoriel de $\mathbb{R}^3$ engendré par le vecteur $v_1 = (1,1,1)$. On note v_2 le vecteur $(1,0,-1)$ et v_3 le vecteur $(0,1,1)$, puis on complète la famille libre $\{v_1\}$ de $\mathbb{R}^3$ en la base $\mathcal{B} = \{v_1, v_2, v_3\}$ (voir également exemple 5.3.4). On considère ensuite la base duale $\mathcal{B}^* = \{v_1^*, v_2^*, v_3^*\}$ et, d'après ce que l'on a vu dans la démonstration précédente, $F^0 = \text{Vect}\,\{v_2^*, v_3^*\}$. L'expression de v_2^* sur $\mathbb{R}^3$ est $v_2^* : \begin{array}{ccc} \mathbb{R}^3 & \to & \mathbb{R} \\ (x_1,x_2,x_3) & \mapsto & x_2 - x_3 \end{array}$ et l'expression de v_3^* sur $\mathbb{R}^3$ est $v_3^* : \begin{array}{ccc} \mathbb{R}^3 & \to & \mathbb{R} \\ (x_1,x_2,x_3) & \mapsto & -x_1 + 2x_2 - x_3 \end{array}$. Ainsi

$$\begin{array}{rcl} F = \left(F^0\right)^0 & = & \left\{(x_1,x_2,x_3) \in \mathbb{R}^3 \mid v_2^*(x_1,x_2,x_3) = 0, v_3^*(x_1,x_2,x_3) = 0\right\} \\ & = & \left\{(x_1,x_2,x_3) \in \mathbb{R}^3 \mid x_2 - x_3 = 0, -x_1 + 2x_2 - x_3 = 0\right\}. \end{array}$$

Une méthode analogue permet d'obtenir, à partir d'une description de F comme ensemble des solutions d'un système d'équations linéaires linéairement indépendantes, une base de F :

Exemple 5.5.5. Notons $\mathcal{B}_0 = \{e_1, e_2, e_3\}$ la base canonique de $\mathbb{R}^3$ et considérons les formes linéaires $\varphi_1 = e_1^* + e_2^* + e_3^*$ et $\varphi_2 = -e_1^* + e_3^*$ sur $\mathbb{R}^3$. On note $W := \text{Vect}\{\varphi_1, \varphi_2\} \subset \mathbb{R}^{3*}$ et on cherche à déterminer une base de $W^0 = \left\{(x_1,x_2,x_3) \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = 0, -x_1 + x_3 = 0\right\}$.

Tout d'abord, remarquons que les formes linéaires φ_1 et φ_2 sont linéairement indépendantes : on peut par exemple constituer la matrice dont les colonnes sont les coordonnées de φ_1 et φ_2 dans la base $\mathcal{B}_0^*$ et montrer qu'elle est bien de rang 2. On note ensuite $\varphi_3 := e_2^* + e_3^*$ et on complète la famille libre $\{\varphi_1, \varphi_2\}$ en la base $\mathcal{C} := \{\varphi_1, \varphi_2\}$ de $\left(\mathbb{R}^3\right)^*$. D'après l'exemple 5.4.3, la base préduale de $\mathcal{C}$ est la base $\mathcal{B} = \{(1,-1,1), (0,-1,1), (-1,2,-1)\}$ de $\mathbb{R}^3$ et, d'après la démonstration de la proposition 5.5.3, $W^0 = \text{Vect}\{(-1,2,-1)\}$.

Remarque 5.5.6. On a $\{0_E\}^0 = E^*$, $E^0 = \{0_{E^*}\}$, $\{0_{E^*}\}^0 = E$ et $(E^*)^0 = \{0_E\}$.

5.6 Application transposée

La dualité linéaire va également nous permettre de donner une interprétation vectorielle à l'opération de transposition sur les matrices.

Soient E et F deux espaces vectoriels sur $\mathbb{K}$.

Définition 5.6.1. *Soit $f \in \mathcal{L}(E,F)$. On appelle transposée de f l'application linéaire*

$${}^t f : \begin{array}{ccc} F^* & \to & E^* \\ \varphi & \mapsto & \varphi \circ f \end{array}$$

Remarque 5.6.2.

- Si $\varphi \in F^* = \mathcal{L}(F, \mathbb{K})$, on a bien $\varphi \circ f \in E^* = \mathcal{L}(E, \mathbb{K})$.
- L'application ${}^t f$ est bien linéaire : si $\varphi, \psi \in F^*$ et $\lambda, \mu \in \mathbb{K}$,

$$\begin{array}{rcl} {}^t f\left(\lambda\varphi + \mu\psi\right) & = & (\lambda\varphi + \mu\psi) \circ f \\ & = & \lambda\varphi \circ f + \mu\psi \circ f \\ & = & \lambda\, {}^t f(\varphi) + \mu\, {}^t f(\psi) \end{array}$$

Les propriétés de base de la transposée d'une application linéaire sont réunies dans la proposition suivante :

Proposition 5.6.3. *1.* $^t(\mathrm{Id}_E) = \mathrm{Id}_{E^*}$.

2. Si $f, g \in \mathcal{L}(E,F)$ *et* $\lambda, \mu \in \mathbb{K}$, $^t(\lambda f + \mu g) = \lambda\, {}^t f + \mu\, {}^t g$.

3. Si G *est un espace vectoriel sur* $\mathbb{K}$, $f \in \mathcal{L}(E,F)$ *et* $g \in \mathcal{L}(F,G)$, $^t(g \circ f) = {}^t f \circ {}^t g$.

4. Si f *est une application linéaire bijective de* E *dans* F, *alors* $^t f : F^* \to E^*$ *est également bijective et* $\left({}^t f\right)^{-1} = {}^t\left(f^{-1}\right)$.

Démonstration. 1. Pour tout $\varphi \in E^*$, on a

$$^t(\mathrm{Id}_E)(\varphi) = \varphi \circ \mathrm{Id}_E = \varphi = \mathrm{Id}_{E*}(\varphi).$$

2. Pour tout $\varphi \in F^*$, on a

$$^t(\lambda f + \mu g)(\varphi) = \varphi \circ (f + g) = \varphi \circ f + \varphi \circ g = {}^t f(\varphi) + {}^t g(\varphi) = \left({}^t f + {}^t g\right)(\varphi).$$

3. Pour tout $\varphi \in G^*$, on a

$$^t(g \circ f)(\varphi) = \varphi \circ (g \circ f) = (\varphi \circ g) \circ f = {}^t g(\varphi) \circ f = {}^t f\left({}^t g(\varphi)\right) = \left({}^t f \circ {}^t g\right)(\varphi).$$

4. On a $^t f \circ {}^t\left(f^{-1}\right) = {}^t\left(f^{-1} \circ f\right) = {}^t(\mathrm{Id}_E) = \mathrm{Id}_{E^*}$ et $^t\left(f^{-1}\right) \circ {}^t f = {}^t\left(f \circ f^{-1}\right) = {}^t(\mathrm{Id}_F) = \mathrm{Id}_{F^*}$.

□

On en vient à la justification matricielle de l'appellation "transposée" et, "réciproquement", à une interprétation vectorielle, via la dualité, de la transposition matricielle :

Proposition 5.6.4. *On suppose que les espaces vectoriels* E *et* F *sont tous deux de dimension finie. Soient alors* $\mathcal{B} = \{e_1, \ldots, e_n\}$ *une base de* E *et* $\mathcal{C} = \{v_1, \ldots, v_m\}$ *une base de* F, *et soit* $f \in \mathcal{L}(E,F)$. *On a*

$$\mathrm{Mat}_{\mathcal{C}^*,\mathcal{B}^*}\left({}^t f\right) = {}^t\,\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(f),$$

autrement dit la matrice de la transposée $^t f$ *de* f *dans les bases duales* $\mathcal{C}^*$ *et* $\mathcal{B}^*$ *est la transposée de la matrice de* f *dans les bases* $\mathcal{B}$ *et* $\mathcal{C}$, *ou encore, symétriquement, la transposée de la matrice de* f *dans les bases* $\mathcal{B}$ *et* $\mathcal{C}$ *est la matrice de la transposée* $^t f$ *de* f *dans les bases duales* $\mathcal{C}^*$ *et* $\mathcal{B}^*$.

Démonstration. Soit $j \in \{1, \ldots, m\}$ et notons $A = (a_{k\,l})_{1 \leqslant k \leqslant m, 1 \leqslant l \leqslant n} := \mathrm{Mat}_{\mathcal{B},\mathcal{C}}(f)$. Alors

$$\begin{aligned} {}^t f\left(v_j^*\right) = \underbrace{v_j^* \circ f}_{\in E^*} &= \sum_{i=1}^n v_j^* \circ f(e_i) e_i^* \text{ăpar la proposition 5.3.6, 1.} \\ &= \sum_{i=1}^n v_j^*\left(f(e_i)\right) e_i^* \\ &= \sum_{i=1}^n v_j^* \left(\sum_{k=1}^m a_{k\,i} v_k\right) e_i^* \\ &= \sum_{i=1}^n \sum_{k=1}^m a_{k\,i} v_j^*(v_k) e_i^* \\ &= \sum_{i=1}^n a_{j\,i} e_i^* \end{aligned}$$

Ainsi la matrice $\mathrm{Mat}_{\mathcal{C}^*,\mathcal{B}^*}\left({}^t f\right)$ est exactement la transposée de la matrice $A = \mathrm{Mat}_{\mathcal{B},\mathcal{C}}(f)$. □

Remarque 5.6.5. Ce résultat est également à mettre en lien avec la propriété 5.4.1 concernant la matrice de passage d'une base duale à une autre : si $\mathcal{B}$ et $\mathcal{B}'$ sont deux bases d'un espace vectoriel de dimension finie E, la matrice de passage $P_{\mathcal{B}\to\mathcal{B}'}$ est la matrice $\mathrm{Mat}_{\mathcal{B}',\mathcal{B}}(\mathrm{Id}_E)$ de l'identité de E dans les bases $\mathcal{B}'$ et $\mathcal{B}$, et la transposée de cette matrice est, d'après les propositions précédentes, la matrice $\mathrm{Mat}_{\mathcal{B}^*,\mathcal{B}'^*}(\mathrm{Id}_{E^*})$, i.e. la matrice de passage $P_{\mathcal{B}'^*\to\mathcal{B}^*}$. Autrement dit, ${}^tP_{\mathcal{B}\to\mathcal{B}'} = P_{\mathcal{B}'^*\to\mathcal{B}^*}$ et donc, de façon équivalente, $P_{\mathcal{B}^*\to\mathcal{B}'^*} = {P_{\mathcal{B}'^*\to\mathcal{B}^*}}^{-1} = {{}^tP_{\mathcal{B}\to\mathcal{B}'}}^{-1} = {}^tP_{\mathcal{B}'\to\mathcal{B}}$.

Ce point de vue "vectoriel" sur la transposition matricielle permet également, à partir de la proposition 5.6.3 et de la correspondance entre produit de matrices et composition d'applications linéaires, de montrer sans calcul les propriétés matricielles "${}^t(AB) = {}^tB\,{}^tA$" et "$\left({}^tA\right)^{-1} = {}^t\left(A^{-1}\right)$".

5.7 Bidual

Soit E un espace vectoriel sur $\mathbb{K}$. Son dual E^* est également un espace vectoriel sur $\mathbb{K}$: on peut donc aussi considérer son dual que l'on note E^{**}. On a alors, par définition,

$$E^{**} = (E^*)^* = \mathcal{L}(E^*, \mathbb{K}) = (\mathcal{L}(E, \mathbb{K}))^* = \mathcal{L}(\mathcal{L}(E, \mathbb{K}), \mathbb{K}).$$

Définition 5.7.1. *On appelle E^{**} le* bidual de E.

On a vu qu'un espace vectoriel de dimension finie est isomorphe à son dual et donc, comme le dual est isomorphe à son propre dual, à son bidual. Néanmoins, comme on l'a dit plus haut, on ne dispose pas, en général, d'isomorphisme "canonique" entre un espace vectoriel de dimension finie et son dual. Une propriété remarquable du bidual est que, en dimension finie, tout espace vectoriel est canoniquement isomorphe à son bidual :

Proposition 5.7.2. *On suppose que E est de dimension finie. Alors E est canoniquement isomorphe à son bidual E^{**}, i.e. on peut construire un isomorphisme linéaire de E sur E^{**} sans faire appel à des choix de bases.*

Démonstration. Pour $v \in E$, définissons tout d'abord l'application $\Phi_v : \begin{array}{ccc} E^* & \to & \mathbb{K} \\ \varphi & \mapsto & \varphi(v) \end{array}$ (l'application d'"évaluation" des formes linéaires sur E en le vecteur v). Pour tout $v \in E$, l'application Φ_v est linéaire : si $\varphi, \psi \in E^*$ et $\lambda, \mu \in \mathbb{K}$,

$$\Phi_v\left(\lambda\varphi + \mu\psi\right) = \left(\lambda\varphi + \mu\psi\right)(v) = \lambda\varphi(v) + \mu\psi(v) = \lambda\Phi_v(\varphi) + \mu\Phi_v(\psi).$$

Ainsi, pour tout $v \in E$, $\Phi_v \in \mathcal{L}\left(E^*, \mathbb{K}\right) = E^{**}$.

On considère alors l'application

$$\Phi : \begin{array}{ccc} E & \to & E^{**} \\ v & \to & \Phi_v \end{array}$$

Φ est une application linéaire : si $v, w \in E$ et $\lambda, \mu \in \mathbb{K}$ alors, pour tout $\varphi \in E^*$,

$$\begin{array}{rcl} \Phi\left(\lambda v + \mu w\right)(\varphi) & = & \Phi_{\lambda v + \mu w}(\varphi) \\ & = & \varphi\left(\lambda v + \mu w\right) \\ & = & \lambda\varphi(v) + \mu\varphi(w) \\ & = & \lambda\Phi_v(\varphi) + \mu\Phi_w(\varphi) \\ & = & \lambda\Phi(v)(\varphi) + \mu\Phi(w)(\varphi) \\ & = & \left(\lambda\Phi(v) + \mu\Phi(w)\right)(\varphi) \end{array}$$

i.e. $\Phi\left(\lambda v + \mu w\right) = \lambda\Phi(v) + \mu\Phi(w)$.

On montre enfin que l'application linéaire $\Phi : E \to E^{**}$ est bijective : comme $\dim\left(E^{**}\right) = \dim\left(E^*\right) = \dim(E)$, on peut se contenter de montrer que Φ est injective. Soit donc $v \in E$ tel que $\Phi(v) = \Phi_v = 0$, i.e., pour tout $\varphi \in E^*$, $\varphi(v) = 0$. En particulier, si l'on considère une base $\mathcal{B} = \{e_1, \ldots, e_n\}$ de E et sa base duale $\mathcal{B}^* = \{e_1^*, \ldots, e_n^*\}$, on a, pour tout $j \in \{1, \ldots, n\}$, $e_j^*(v) = 0$ et donc $v = \sum_{j=1}^{n} e_j^*(v)e_j = 0$: Φ est donc bien injective.

Ainsi, l'application Φ est bien un isomorphisme linéaire de E sur E^{**} (et sa définition ne dépend pas d'un choix de bases). □